*e-X*tra*

REVIEW

Roadmap to New Virulence Determinants in *Pseudomonas syringae*: Insights from Comparative Genomics and Genome Organization

Magdalen Lindeberg,[1] Christopher R. Myers,[2] Alan Collmer,[1] and David J. Schneider[1,3]

[1]Department of Plant Pathology, Cornell University, Ithaca, NY 14853, U.S.A.; [2]Computational Biology Service Unit, Center for Advanced Computing, Cornell University, Ithaca, NY 14853, U.S.A.; [3]U. S. Department of Agriculture Agricultural Research Service, Ithaca, NY 14853, U.S.A.

Systematic comparison of the current repertoire of virulence-associated genes for three *Pseudomonas syringae* strains with complete genome sequences, *P. syringae* pv. *tomato* DC3000, *P. syringae* pv. *phaseolicola* 1448A, and *P. syringae* pv. *syringae* B728a, is prompted by recent advances in virulence factor identification in *P. syringae* and other bacteria. Among these are genes linked to epiphytic fitness, plant- and insect-active toxins, secretion pathways, and virulence regulators, all reflected in the recently updated DC3000 genome annotation. Distribution of virulence genes in relation to *P. syringae* genome organization was analyzed to distinguish patterns of conservation among genomes and association between genes and mobile genetic elements. Variable regions were identified on the basis of deviation in sequence composition and gaps in syntenic alignment among the three genomes. Mapping gene location relative to the genome structure revealed strong segregation of the HrpL regulon with variable genome regions (VR), divergent distribution patterns for toxin genes depending on association with plant or insect pathogenesis, and patterns of distribution for other virulence genes that highlight potential sources of strain-to-strain differences in host interaction. Distribution of VR among other sequenced bacterial genomes was analyzed and future plans for characterization of this potential reservoir of virulence genes are discussed.

Additional keywords: Arabidopsis, nonribosomal peptide synthase, pathogenicity island, REP sequence.

The genus *Pseudomonas* occupies a wide range of environments, with many strains and species adapted to niches associated with animals, plants, and insects. To date, genome sequences have been completed for 13 pseudomonads, with three sequences derived from isolates of *Pseudomonas syringae*, a species characterized by its ability to epiphytically colonize and cause disease in plants. Many of the individual *P. syringae* pathovars, named for the plant species from which they were isolated, share a generalized disease cycle characterized by dissemination in association with plant material, colonization of the plant surface, and onset of disease upon entry into the plant apoplast. However, dimensions of the bacterial-host interaction, including epiphytic fitness, host range, characteristic symptoms, and disease severity, vary substantially among pathovars and strains.

The three sequenced *P. syringae* strains represent each of the major monophyletic groups within the species (Sarkar and Guttman 2004; Sawada et al. 2002) and differ in the nature of their interaction with host plants (Buell et al. 2003; Feil et al. 2005; Joardar et al. 2005a). *P. syringae* pv. *tomato* DC3000 is a causal agent of bacterial speck on tomato and *Arabidopsis*. Effector proteins secreted by the type III secretion system (T3SS) have been extensively characterized in this strain, and it has emerged as an important model system for exploring interactions between bacterial and plant gene products during the infection process (Lindeberg et al. 2006; Nomura et al. 2005). *P. syringae* pv. *phaseolicola* 1448A is an economically important pathogen of bean and is closely related to strains causing devastating outbreaks in Africa (Fourie 1998). Type III effectors and other members of the HrpL regulon have also been extensively characterized in this strain (Chang et al. 2005; Vencato et al. 2006). *P. syringae* pv. *syringae* B728a is the causal agent of brown spot of bean and, of the three sequenced strains, is best adapted for epiphytic survival and growth on the plant surface (Hirano and Upper 2000; Marco et al. 2005; Monier and Lindow 2003). Though many gene products have been implicated in different stages of host interaction, including T3SS effector proteins, toxins, siderophores, adhesins, and proteins contributing to stress tolerance, knowledge of the full range of factors contributing to host association remains incomplete. What is our present understanding of gene products implicated in adaptation to the plant-associated niche? Of the factors thus far identified, which ones underlie strategies shared by diverse pathovars and which ones confer pathovar-specific phenotypes? These questions can be addressed by comparison of the repertoire of virulence and host-association genes among the three strains, followed by analysis of their locations relative to overall genome structure. Though the ultimate goal of assessing patterns in gene distribution is the identification of regions likely enriched for additional host-association genes, the pathovar-specific and species-specific genes

and regions identified during these analyses represent a valuable pool of potential molecular markers for use in genome-based diagnostics of *P. syringae* isolates.

To date, a systematic three-way comparison has not been conducted for the *P. syringae* pv. *tomato* DC3000, *P. syringae* pv. *phaseolicola* 1448A, and *P. syringae* pv. *syringae* B728a genomes. "Lineage-specific regions" in DC3000 relative to *P. putida* and *P. aeruginosa* were previously identified based on enrichment for mobile genetic elements and DC3000-specific genes (Joardar et al. 2005b), but differences in gene repertoire and organization among *P. syringae* pathovars have not been thoroughly investigated.

Review of the repertoire of genes implicated in virulence and host association of *P. syringe* is timely, given the significant advances in identification and confirmation of predicted virulence factors since publication of the original genome annotations (Buell et al. 2003; Feil et al. 2005; Joardar et al. 2005a). Some of this work has been conducted in *P. syringae*, while other insights are derived from ongoing research in other bacterial pathogens that has greatly expanded the pool of characterized host-association factors with which the *P. syringae* genomes can be screened. Gene products with confirmed and candidate roles in virulence and host association are described in the following sections, with particular emphasis on those more recently characterized. These new findings are being incorporated into the appropriate genome annotation records at National Center for Biotechnology Information (NCBI) with updates summarized on the Pseudomonas-Plant Interaction website.

Recent advances in characterization of *P. syringae* virulence and host-association factors.

Epiphytic fitness determinants—biofilms, *osmoprotection*, *and iron uptake.* Survival on the leaf surface prior to or exclusive of subsequent infection requires the ability to withstand stresses specific to the epiphytic environment and acquire nutrients in this highly circumscribed environment (Hirano and Upper 2000). Production of type IV pili and the exopolysaccharide alginate, phenotypes associated with attachment and biofilm-associated protection in numerous genera, have been previously shown to contribute to either epiphytic fitness, virulence, or both in *P. syringae* (Chang et al. 2007; Roine et al. 1998; Yu et al. 1999). However, neither alginate nor the polysaccharide levan appear to be required for biofilm formation in vitro, suggesting that as-yet-uncharacterized polysaccharides are critical to *P. syringae* biofilm formation (Laue et al. 2006). A recently identified candidate is the mannose-rich Psl (polysaccharide synthesis locus) exopolysaccharide, shown by Ma and associates (2006) to play a critical role in adhesion initiation and biofilm maintenance by *P. aeruginosa*. Given that orthologs of all 11 loci essential for Psl biosynthesis are present in *P. syringae* (PSPTO_3529 to 3539), this polysaccharide may play a significant role in *P. syringae*–host interactions. A second polysaccharide with potential involvement in *P. syringae* biofilm formation is acetylated cellulose, implicated in biofilm formation and epiphytic fitness in *P. fluorescens* SBW25 (Gal et al. 2003; Spiers et al. 2003). The complete *wss* biosynthetic cluster, characterized in SBW25, is present in *P. syringae* pv. *tomato* DC3000 (PSPTO_1026 to 1034) and production of cellulose by DC3000 has been confirmed (Ude et al. 2006).

Bacteria living epiphytically experience significant oxidative and osmotic stress (Hirano and Upper 2000). Tolerance of UV and oxidative stress is typically addressed through deployment of enzymes involved in quenching of reactive oxygen species as well as *rulAB*-encoded DNA polymerase for repair of UV damage (Sundin and Murillo 1999). Osmotic stress is mitigated by exopolysaccharide production and import of osmoprotective compounds that, like betaine, function directly as compatible solutes or, like choline, serve as precursors for compatible solute synthesis (Boncompagni et al. 1999; Landfald and Strom 1986; Wood 1999). Recent investigation of osmoprotection mechanisms in *P. syringae* pv. *tomato* DC3000 has led to identification of two transporters shown to be capable of in vitro uptake of osmoprotective compounds: i) the OpuC ABC transporter (PSPTO_4575 to 4578), implicated in the uptake of betaine and choline (Chen and Beattie 2007) and ii) the BCCT family transporter, BetT (PSPTO_5269), involved in choline transport (Chen and Beattie 2008). Interestingly, uptake of exogenous choline confers greater osmoprotection in DC3000 than betaine, the typical substrate observed for other bacteria (Miller and Wood 1996). Conversion of choline to betaine is likely accomplished by gene products encoded by the *betIBA* locus (PSPTO_0440 to 0441, 0443), similar to those characterized in *Escherichia coli* (Lamark et al. 1991).

Acquisition of nutrients is similarly essential for survival on the leaf surface, with uptake of limited iron accomplished by deployment of iron-chelating siderophores. The most definitive evidence supporting the role of siderophores in epiphytic growth comes from studies showing a growth defect for siderophore-deficient *P. syringae* pv. *syringae* B728a in the presence of plant-derived iron-chelated compounds (Karamanoli and Lindow 2006). In planta expression of the siderophore pyoverdin has been recently confirmed in *P. syringae* pv. *tomato* DC3000 grown in iron-limiting conditions (Jones et al. 2007), and a fitness benefit was observed when expressed under iron-limiting conditions (M. W. Wildermuth, *personal communication*). In contrast, the siderophore yersiniabactin, though clearly expressed, does not appear to contribute positively to fitness of B728a under iron-limiting conditions (Jones et al. 2007).

The three *P. syringae* genome sequences also contain numerous genes annotated as TonB-dependent receptors (TBDR). Though generally implicated in uptake of siderophores, recent research in *Xanthomonas campestris* has revealed that TBDR-mediated uptake of carbohydrates is required for full virulence on *Arabidopsis*, suggesting that selected TBDR orthologs in *P. syringae* may also play determining roles in the uptake of critical, nonferric substrates (Blanvillain et al. 2007).

T3SS effector proteins, *helpers*, *and other members of the HrpL regulon.* Following passage from the epiphytic surface through stomatal openings, bacteria must actively defeat plant defense responses in order to progress to the disease state. As agents deployed to suppress both basal and gene-for-gene resistance following bacterial entry into the plant, Hop effector proteins represent a critically important class of virulence factors (Nomura et al. 2005). Characterized on the basis of their regulation by the HrpL alternate sigma factor and translocation by the T3SS, 46 families of Hop effectors and seven families of so-called "helper" proteins involved in facilitating transit of the effector proteins through the T3SS have been identified in the sequenced *P. syringae* strains (Chang et al. 2005; Ferreira et al. 2006; Fouts et al. 2002; Greenberg and Vinatzer 2003; Guttman et al. 2002; Petnicki-Ocwieja et al. 2002; Vencato et al. 2006) and have been reviewed previously (Lindeberg et al. 2006). HrpL-dependent expression of selected non-*hop* genes has also been experimentally confirmed in *P. syringae* pv. *tomato* DC3000 and *P. syringae* pv. *phaseolicola* 1448A (Ferreira et al. 2006; Vencato et al. 2006). Ongoing studies have revealed a wide variety of in planta target sites and specific activities for individual effectors including AvrPto1 (Lin and Martin 2007; Xing et al. 2007), AvrPtoB (Abramovitch and Martin 2005; Rosebrock et al. 2007), AvrE1 (Badel et al. 2006), HopF2 (Robert-Seilaniantz et al. 2006), HopI1 (Jelenska et al. 2007), HopU1 (Fu et al. 2007), HopM1 (Badel et al. 2006),

HopN1 (Lopez-Solanilla et al. 2004), HopQ1-1 (Robert-Seilaniantz et al. 2006; Wei et al. 2007), and HopAO1 (Bretz et al. 2003; Espinosa et al. 2003).

Recent investigations also shed light on the respective contributions of T3SS helper proteins to the translocation process. Kvitko and associates (2007) report that HopAK1 and HopP1 share essential harpin-like properties with the previously identified HrpZ and HrpW and that HrpW, HrpZ, and HopAK1 play essential roles in translocation of effector proteins into the plant cell. These three proteins also appear to have an overlapping function with the predicted translocator, HrpK (Kvitko et al. 2007; Petnicki-Ocwieja et al. 2005). The class of helper proteins with lytic transglycosylase activity, facilitating passage of the Type III pilus component through the bacterial peptidoglycan layer, includes HrpH and HopAJ1 and the harpin-like HopP1. HrpH appears to be the primary specialized lytic transglycosylase in *P. syringae*, with HopAJ1 and HopP1 having secondary, overlapping functions (Oh et al. 2007).

Noneffector members of the HrpL regulon linked to a role in either fitness, virulence, or both include an ApbE ortholog (PSPTO_2105) implicated in the alternative pyrimidine biosynthetic pathway in *Salmonella typhimurium* (Beck and Downs 1998) and an alcohol dehydrogenase (PSPTO_0834). Both loci have been shown to impact bacterial growth in *Arabidopsis* (Vencato et al. 2006). Though separated from an upstream HrpL binding site by an intervening insertion sequence, a potential role for the DC3000 gene cluster PSPTO_0873 to 0875 is suggested by the finding that the orthologous *hsv* operon in *Erwinia amylovora* contributes to virulence on apple shoots (Oh et al. 2005).

Toxins and other T3SS-independent virulence factors. P. syringae strains typically produce phytotoxins using nonribosomal peptide synthases and polyketide synthases, which contribute significantly to virulence in the pathovars in which they occur (Bender et al. 1999; Donadio et al. 2007; Raaijmakers et al. 2006). Previously characterized toxins include lipodepsipeptidic toxins such as syringomycin and syringopeptin, involved in both attachment and surfactant activity (Raaijmakers et al. 2006), and antimetabolite toxins such as phaseolotoxin and tabtoxin, which target specific metabolic pathways in the host. *P. syringae* pv. *tomato* DC3000 produces coronatine, an adduct of the polyketide coronafacic acid and an isoleucine derivative, coronamic acid. In addition to roles in suppression of salicylic acid-mediated defense (Kloek et al. 2001) and activation of the jasmonic acid (JA)-signaling pathway through mimicry of JA (Zhao et al. 2003), recent research has revealed a critical role for coronatine in suppression of stomatal defense (Melotto et al. 2006; Underwood et al. 2007).

Biosynthetic loci for two additional *P. syringae* toxins have recently been identified. The first of these, characterized in DC3000, consists of *syfRABCD* (PSPTO_2828 to 2832), which encodes gene products responsible for production of six related lipodepsipeptidic compounds termed "syringofactins" SyrA to SyrF (Berti et al. 2007). Though linear in structure, the syringofactins exhibit properties similar to previously characterized cyclic lipodepsipeptidic toxins and are required for both surfactant activity and swarming motility (Berti et al. 2007).

The second new toxin is the ornithine *N*-acetyltransferase inhibitor "mangotoxin," identified in *P. syringae* pv. *syringae* UMAF0158A as a virulence factor in apical necrosis of mango (Arrebola et al. 2003). Gene clusters orthologous to the *P. syringae* pv. *syringae* UMAF0158A mangotoxin biosynthetic genes are found in other *P. syringae* genomes (e.g., PSPTO_5452 to 5458 in DC3000). A potentially widespread role in virulence is predicted by the observation that UMAF0158A impaired in mangotoxin production is reduced in virulence on tomato (Arrebola et al. 2007).

Screening the *P. syringae* genomes for genes orthologous to virulence determinants in diverse bacterial pathogens has revealed a growing list of candidate toxins and other gene products with surprising similarity to virulence determinants deployed by bacterial pathogens of insects. For example, the two most significant virulence factors thus far characterized in *Pseudomonas entomophila* L48, a recently identified and sequenced pathogen of genus *Drosophila*, are a nonribosomal peptide synthase and the metalloprotease AprA. Encoded by PSEEN_0132, the nonribosomal peptide synthase and flanking loci are highly orthologous to none other than the biosynthetic cluster responsible for mangotoxin production (I. Vallet-Gely, *personal communication*). The metalloprotease AprA, abundantly secreted and required for *P. entomophila* survival and pathogenicity in host *Drosophila* (Liehl et al. 2006), also has a closely related ortholog in *P. syringae* (PSPTO_3332) (Vodovar et al. 2006).

Other genes present in the *P. syringae* genomes share similarity with insect active toxins deployed by nematode symbionts *Xenophabdus nematophila* and *Photorhabdus luminescens*. The binary toxin *xaxAB*, encoded by *X. nematophila*, was initially reported to have a single orthologous locus in *P. syringae* pv. *syringae* B728a (Vigneux et al. 2007), but paralogous loci are actually present at two locations in all three *P. syringae* genomes (PSPTO_4570 to 4571 and PSPTO_5625 and 4287 in DC3000). Orthologs of toxin complex (Tc) genes, first characterized in *Photorhabdus luminescens* and encoding high molecular weight proteins with insecticidal activity (Pinheiro and Ellar 2007; Waterfield et al. 2001), are also found at diverse locations in the *P. syringae* genomes, with the cluster corresponding to *P. syringae* pv. *tomato* DC3000 loci PSPTO_4341 to 4344 encoding representatives of each of the three major families of Tc proteins (TcdA, TcdB, TccC).

Though experimental evidence is less definitive, chitinases, hemolysins, and lipases have also been implicated in bacterial pathogenesis of insects (Vodovar et al. 2006). Predicted members of these gene families are found in one or more of the *P. syringae* genomes.

Secretion pathways. All three *P. syringae* genomes possess a well-characterized T3SS, annotated orthologs of the Type II secretion components, and numerous Type I ABC transporters. Additional cryptic copies of the type II and III pathways are found in *P. syringae* pv. *phaseolicola* 1448A. A more recently identified *P. syringae* secretion pathway is the twin-arginine transport (Tat) pathway, experimentally characterized in *P. syringae* pv. *tomato* DC3000 (PSPTO_5155 to 5157) and involved in translocation of folded-proteins protein complexes across the cytoplasmic membrane prior to type II-dependent passage across the outer membrane (Maillard et al. 2007). Disruption of the Tat pathway in DC3000 results in a reduction in virulence on *Arabidopsis*, with two lipases, PlcA1 (PSPTO_3648) and PlcA2 (PSPTO_B0005), shown to employ the pathway for extracellular localization (Bronstein et al. 2005).

Genes encoding components of the newly described type VI secretion system (T6SS) are also present at two locations in *P. syringae* pv. *tomato* DC3000. Mutations in the T6SS significantly reduce virulence in several bacterial pathogens, including *Vibrio cholerae*, *P. aeruginosa*, and *Burkholderia pseudomallei* (Mougous et al. 2006; Pukatzki et al. 2006; Schell et al. 2007). Recent work suggests that elements of the secretion apparatus itself assemble into a structure that punctures cells in a manner analogous to phage tail spikes, but the substrates delivered through this structure are as yet uncharacterized (Pukatzki et al. 2007).

Regulation of virulence-related genes. Recognition of environmental variables and coordinated control of the gene products required for successful exploitation of and transition be-

tween the epiphytic and apoplastic niches is accomplished by a complex network of regulatory controls, recently reviewed by Mole and associates (2007) for diverse phytopathogenic bacteria. Among these controls are an array of alternate sigma factors and quorum-sensing signals, the latter being linked to regulation of traits associated with epiphytic fitness and virulence in *P. syringae* pv. *syringae* B728a (Quinones et al. 2005). Research on factors controlling expression of the acyl homoserine lactone signal (AHL) integral to quorum sensing has led to identification of two new regulators, AefR and PsrA. The first of these upregulates AHL synthesis in *P. syringae* pv. *tomato* DC3000 and *P. syringae* pv. *syringae* B728a, while PsrA downregulates AHL synthesis through its effect on both AefR and RpoS (Chatterjee et al. 2007; Quinones et al. 2004). The recently characterized LadS activator and RetS repressor (Ventre et al. 2006), integral to switching *P. aeruginosa* from its free-living motile state to multicell biofilm via control of small RNA transcription, have not been studied in *P. syringae*, but orthologs of *retS* and *rsmA* encoding a small RNA binding protein are present, suggesting that a related network may contribute to regulation of fitness and virulence determinants in *P. syringae*.

As previously mentioned, the T3SS and substrates belong to a group of genes regulated by the alternative sigma factor HrpL, expression of which is, in turn, regulated by NtrC-like transcription activators HrpR and HrpS (Chatterjee et al. 2002). Genes regulated by HrpL have been comprehensively characterized in *P. syringae* pv. *tomato* DC3000 (Fereirra et al. 2006; Fouts et al. 2002) and *P. syringae* pv. *phaseolicola* 1448A (Chang et al. 2005; Vencato et al. 2006) and those regulated by HrpRS in DC3000 (Lan et al. 2006). HrpL binding sites have since been incorporated into the *P. syringae* genome annotations and represent a valuable addition to information stored in those records.

In view of the expanding repertoire of gene products implicated in bacterial-host interaction, the following questions arise. Which among these gene products are limited to specific lineages, indicative of a role in pathovar or strain-specific adaptations, and which are more widely distributed, suggesting involvement in more universal functions shared by multiple pathovars in diverse plant hosts? Are the widely distributed factors found in syntenically conserved regions consistent with ancient association with the lineage, or are they found in regions exhibiting a high degree of genetic exchange, indicative of recent acquisition or rearrangement? For genes subject to high rates of duplication and horizontal transfer, is there evidence of persistent linkages between virulence genes or with mobile genetic elements that could reveal functional associations or underlying mechanisms of transfer?

Mapping core and variable genome regions (VR) in the *P. syringae* genomes.

Bacterial genomes are composed of mosaics of relatively stable core regions interspersed with regions exhibiting either high levels of variation in content, alignment, or both relative to related genomes (Hacker and Kaper 2000; Welch et al. 2002). Sometimes described as genomic islands or pathogenicity islands (when components are implicated in virulence-related functions), these regions are shaped by a range of processes including excision, recombination, duplication, and horizontal transfer (Gal-Mor and Finlay 2006). In addition to alterations in alignment or gene synteny relative to related genomes, VR are often characterized by deviations in the composition of the DNA sequence and typically are associated with features such as mobile genetic elements, tRNAs, and intergene repetitive elements (REP sequences) (Karlin 2001; Lawrence and Ochman 1997; Paulsen et al. 2003; Tobes and Pareja 2005).

Though VR are frequently identified on the basis of enrichment for genes unique to a particular genome or group of closely related genomes, the approach described here focuses on gaps in the genome alignment (syntenic divergence) to maximize identification of unique genes in addition to those present in multiple genomes but acquired by horizontal transfer or subject to intragenomic rearrangement. Sites of syntenic divergence were identified by generating pairwise comparisons of DC3000, 1448A, and B728a using BLASTN (Altschul et al. 1990) and MSPcrunch (Sonnhammer and Durbin 1994) and by analysis of the resulting output with the Artemis Comparison Tool (Carver et al. 2005).

Compositional deviation in the DNA sequence was also measured not only to evaluate the likelihood of horizontal acquisition of syntenically disrupted regions but also to reveal regions of more ancient horizontal transfer in which syntenic disruption may not be present. Deviation in sequence composition is frequently quantitated using sequence properties such as variation in GC content and skews in the dinucleotide or trinucleotide composition (Hsiao et al. 2005; Paulsen et al. 2003); however, the fixed-order methods typically used do not maximally capture local composition. The Alien_hunter application compensates by simultaneously assessing compositional deviation at multiple levels, using variable order compositional distribution, with a two-state second-order hidden Markov model applied to optimize boundary prediction (Vernikos and Parkhill 2006). Previously shown to identify novel pathogenicity islands in *Salmonella enterica* serovar Typhi CT18 (Vernikos and Parkhill 2006), Alien_hunter was used to analyze the genomes of the three fully sequenced *P. syringae* pathovars, using the default threshold for compositional deviation.

Synthesizing data derived from the two approaches, VR were defined as those exhibiting either gaps in genome alignment ≥5 kb in at least one pairwise comparison or, in the absence of syntenic divergence, deviation in sequence composition predicted by Alien_hunter at corresponding locations in all three strains. Boundaries of syntenically divergent regions were established based on the locations of flanking regions exhibiting >5-kb alignment for all three genomes. For VR identified by both syntenic disruption and compositional deviation, the final region size was determined according to the larger of the two predictions. For the purpose of this analysis, the plasmids of *P. syringae* pv. *tomato* DC3000 and *P. syringae* pv. *phaseolicola* 1448A are each considered discrete VR.

As previously described, the *P. syringae* genomes contain several major genome rearrangements relative to one another, with regions as large as 500 kb inverted and rearranged relative to the rest of the genome (Feil et al. 2005; Joardar et al. 2005a). Given that boundaries for individual VR are defined relative to the DNA immediately flanking, placement in relation to the major rearrangements was not an issue for most of the VR identified. However, for those VR located at the boundaries of these rearrangements, region size was determined according to the distance between flanking blocks of aligned DNA, even when the respective flanks were not aligned relative to each other.

Application of these guidelines resulted in the prediction of 107, 87, and 78 VR for *P. syringae* pv. *tomato* DC3000, *P. syringae* pv. *phaseolicola* 1448A, and *P. syringae* pv. *syringae* B728a constituting 36, 25, and 29% of the total genome size, respectively (Table 1). These percentages are higher than the levels of variability determined from pairwise BLASTP analysis of individual open reading frames (ORF) (Feil et al. 2005; Joardar et al. 2005b), consistent with the fact that VR defined here encompass both unique regions and conserved genes either lacking syntenic conservation or conserved and acquired during ancient horizontal transfer. Of the 121 VR identified

among the three genomes, 109 were identified on the basis of both compositional deviation and syntenic disruption, 11 by Alien _hunter alone and eight by syntenic disruption alone. The high level of correspondence between locations of compositional deviation and gaps in genome alignment supports the utility of Alien_hunter and related statistical methods as tools for automated identification of putative VR.

As shown in Table 1, VR can be classified on the basis of their representation in the three genomes, where "unique" corresponds to regions found only in a single genome, "conserved" corresponds to those conserved in content but exhibiting deviation in sequence composition as determined by Alien_hunter, and "hotspot" refers to regions exhibiting either divergent content or compositional deviation or both in all three genomes. "Partial" indicates that a given region exhibits syntenic divergence or compositional deviation or both in two of the three genomes, and "boundary" refers to regions for which synteny cannot be established owing to its location at a site of genome rearrangement. Consistent with its overall larger genome size, DC3000 has both the largest number of VR and the highest percentage of its genome exhibiting variable content, reflected chiefly in a greater number of unique insertions and larger insertion sizes at conserved "hotspots" (Table 1). In contrast, 1448A has only one unique insertion (VR 32) and B728a none, though B728a contains two sizable insertions in boundary regions (VR 68 and 90), accounting for most of its variable genome in this category. The three genomes share eleven conserved regions.

It has been previously demonstrated that REP sequences preferentially segregate with the syntenically conserved core genome (Paulsen et al. 2005) and represent hotspots for genomic rearrangement (Tobes and Pareja 2006). To evaluate placement of REP sequences relative to core and VR, a hidden Markov model based on sequences described by Tobes and Pareja (2005) was constructed and used to search for REP sequences in the *P. syringae* pv. *tomato* DC3000 chromosome (Tobes and Pareja 2005). Of the 368 predicted REP sequences, 89% segregated with the predicted core regions, providing additional support for the core genome predictions in DC3000. Locations of REP sequences have been recently added to the DC3000 genome annotation at NCBI.

It should be noted that variables contributing to compositional variation have a tendency to ameliorate over time, such that more ancient horizontal transfer events are detected with less sensitivity (Lawrence and Ochman 1997). Regions of compositional deviation identified by Alien_hunter may also reflect sequence biases associated with function such as the high inci-dence of GC-rich hydrophobic amino-acid codons associated with membrane proteins. Furthermore, it should neither be assumed that the predicted flexible regions have arisen from single insertions nor that the boundaries given necessarily correspond to the locations of discrete insertion events. As is evident upon examination of the larger flexible regions, many of these regions are the result of multiple overlapping insertions, disruptions, and duplications. Systematic identification of regions transferred during independent events has proven extremely difficult.

Locations of virulence and host-association genes relative to the core and flexible genome.

A map of gene locations and corresponding locus tags relative to the numbered VR is shown in Table 2. *P. syringae* pv. *tomato* DC3000 "*hop* islands," described in previous deletion analyses, are also indicated (Wei et al. 2007). As a general observation, regulatory factors preferentially segregate with the core genome, including the recently described two-component regulatory system clusters, orphaned components, and hybrids, of which 82 of 107 are found in the core regions, and an additional 14 syntenically conserved within VR (Lavin et al. 2007). In contrast, members of the HrpL regulon segregate almost exclusively with VR. Most other classes of virulence and host-association factors are distributed more evenly between the core and variable genomes. Analysis of gene location relative to the core and VR is described for various classes of virulence factors in the following sections, together with insights on the relative timing of gene acquisition and potentially significant linkages among loci.

The HrpL regulon is localized to VR with evidence for comobilization by some members. Members of the HrpL regulon represent the class of virulence genes most strongly segregating with the VR, although, when shared by different strains, their locations are generally syntenic. This phenomenon is observed for the *hrp/hrc* cluster encoding the T3SS, previously shown to be of ancient association with the *P. syringae* lineage (Alfano et al. 2000) and identified here as variable on the basis of compositional deviation (VR 36 and 37). The conserved effector locus (*avrE*, *hopM1*, and *hopAA1*) is also of ancient association but was not identified by Alien_hunter, possibly owing to a greater degree of sequence amelioration.

Of the remaining *hop* genes present in all three strains, *hopAF1*, *I1*, *AH2*, and *AN1* are found in small regions of syntenic conservation within otherwise VR hotspots. Likewise, *hopAS1*, *AE1*, *AB1*, *AJ1*, *AT1*, *AG1*, *AH2*, and *AI1*, shared by two of the three strains, are found at corresponding locations

Table 1. Number and size of variable regions identified by analysis of compositional deviation and syntenic disruption of *Pseudomonas syringae* pv. *tomato* DC3000, *P. syringae* pv. *phaseolicola* 1448A, and *P. syringae* pv. *syringae* B728a genomes [y]

Variable region type	*P. syringae* pv. *tomato* DC3000		*P. syringae* pv. *phaseolicola* 1448A		*P. syringae* pv. *syringae* B728a	
	Number	Size (kb)	Number	Size (kb)	Number	Size (kb)
Total (including plasmids)	105(107)	2,199 (2340)	85(87)	1,562 (1,746)	78	1,751
Unique (including plasmids)	19(21)	271 (412)	1(3)	17 (201)	0	0
Conserved[z]	11	93	11	93	11	95
Hotspot	45	1,237	45	763	45	864
Partial	18	173	19	297	10	128
Boundary	12	425	10	392	12	664

[y] Unique regions are those with an insertion present in one genome but no compositional deviation and gaps in syntenic alignment or insertion of mobile DNA at the corresponding location in the other two genomes. Conserved regions are identified by Alien_hunter as having compositional deviation in all three genomes but without gaps in the syntenic alignment. Hotspot regions are those with gaps in syntenic alignment, compositional deviation, or mobile DNA, singly or in combination, at corresponding locations in all three genomes. The partial designation reflects regions with gaps in syntenic alignment, compositional deviation, or mobile DNA, singly or in combination, at corresponding locations in two of the three genomes. The boundary designation indicates regions located at the boundaries of the major genome rearrangements in which genomic material does not align with either flank, but relative synteny cannot be established in one or more pairwise alignments owing to the rearrangement.

[z] The 2-kb difference in region size observed for the conserved regions in *P. syringae* pv. *syringae* B728a relative to *P. syringae* pv. *tomato* DC3000 and *P. syringae* pv. *phaseolicola* 1448A reflects a discrepancy in the predictions made by Alien_hunter.

Table 2. Locations of virulence and other host-association factors relative to the variable genome regions (VR), listed in the order they occur in the *Pseudomonas syringae* pv. *tomato* DC3000 genome, and identified by analysis of compositional deviation and sequence alignment of the *P. syringae* pv. *tomato* DC3000, *P. syringae* pv. *phaseolicola* 1448A, and *P. syringae* pv. *syringae* B728a genomes

VR[y]	*P. syringae* pv. *syringae* B728a[x]		*P. syringae* pv. *tomato* DC3000		*P. syringae* pv. *phaseolicola* 1448A		*hop* islands[z]
	Gene	Locus	Gene	Locus	Gene	Locus	
1			**hopK1**	**0044**			
			hopY1	**0061**			
3-B					**hopR1**	**0171**	
					Hypothet.	**5226**	
	algZ, R, R3, Q	0054, 0056, 0063 to 0064	algZ, R, R3, Q	0126 to 0127, 0134, 0136	algR3, Q, Z, R	0059, 0061, 0069 to 0070	
10			**MATE efflux**	**0370**			
			Indole 3-acetic acid ligase	**0371**			
11	Chitinase	4777	Chitinase	0400	Chitinase	4807	
	Betaine biosynth.	4732 to 4734	Betaine biosynth.	0440 to 0441, 0443	Betaine biosynth.	4766 to 4768	
13			**hopAS1'**	**0474**	**hopAS1**	**4736**	
14			**hopU1**	**0501**			I
			hopF2, shcF	**0502** to 0503			
15-B	**hopAB1**	**4659**					
	NRPS	4662					
16-B					NRPS	4696	
18			**hopH1**	**0588**			II
			hopC1	**0589**			
19	katA catalase	4522			katA catalase	4565	
	Xylanase	4508			Xylanase	4536	
					NRPS	4546 to 4547, 4750	
					FHA	4560	
20					TBDR	4499	
22			TBDR	0784			
23	Type IV pili biogen.	0714 to 0720	Type IV pili biogen.	0810 to 0816	Type IV pili biogen.	0725 to 0731	
24	rulAB	0735 to 0736	**Oxidoreductase, ribD, hydrolase**	**0834** to 0836	Type IV pilin	0759 to 0760	
	avrRpm1	**0738**					
			Hypothet.	**0851**	**Hypothet.**	**0762**	
			hopAJ1'	**0852**	**hopAJ1'**	**0763**	III
			hopAT1'	**5618**	**hopAT1'**	**05225**	
	Levansucrase	0754	**Transposon (deg.)**	**5617**	**hopG1**	**0767**	
			hsv-like operon	0873 to 0875	Fimbrial proteins	0773 to 0774	
			hopD1	**0876**			
			hopQ1-1	**0877**			IV
			hopR1	**0883**			
					Hemagglutinin	0781	
					avrB4-2	**0784**	
	hopAG1'	**0778**	**hopAG1'**	**0901**			V
	hopAH1	**0779**	**hopAH1**	**0905**			
	hopAI1	**0785**	**hopAI1**	**0906**			
25					Hemolysin	0815	
	pilBCD, pre-pilin	0796 to 0799	pilBCD, pre-pilin	0924 to 0927	pilBCD, pre-pilin	0818 to 0821	
26	Pectate lyase	0852			Pectate lyase	0883	
27			**hopAM1-1**	**1022**			
			wss operon	1026 to 1034			
30			TBDR	1206			
	TBDR	1038	TBDR	1207	TBDR	1091	
31	Tc toxin-TccC	1051	Tc toxin-TccC', TccC	1225, 1231			
	Alginate biosynth.	1052 to 1063	Alginate biosynth.	1232 to 1243	Alginate biosynth.	1107 to 1118	
35			**hopN1**	**1370**			

Continued on following page

[x] Names and locus tags of HrpL-regulated genes are indicated in bold type. biogen. = biogenesis, biosynth. = biosynthesis, deg. = degenerate, EPS = exopolysaccharides, FHA = filamentous hemagglutinin, hypothet. = hypothetical, NRPS = nonribosomal peptide synthase, polysacch. = polysaccharide, put. = putative, TBDR = TonB-dependent receptor, resist. = resistance, T2SS and T3SS, respectively = type II and III secretion systems.

[y] Variable genome regions are listed in the order in which they occur relative to the *P. syringae* pv. *tomato* DC3000 genome. Due to inversions and genome rearrangements, relative ordering of variable regions in *P. syringae* pv. *phaseolicola* 1448A and *P. syringae* pv. *syringae* B728a should not be inferred from the numbers assigned. Individual cells are colored as follows: gold, syntenically conserved host-association factors within the variable genome regions; white, nonsyntenically conserved host-association factors within the variable genome regions; light gray, regions of divergent content within variable genome regions; blue, selected examples of syntenically conserved host-association factors in core genome regions; black, no corresponding variable region is present in the genome, due either to the absence of noncore material or because location at a genome rearrangement boundary prevents evaluation of syntenic correspondence. Locations of regions with similarity to integrative and conjugative elements (ICElands) are outlined in red. Variable regions located at the boundaries of the major genome rearrangements in one or more pairwise comparisons are indicated with a "B". Placement of the "B" variable regions in *P. syringae* pv. *phaseolicola* 1448A and *P. syringae* pv. *syringae* B728a relative to the *P. syringae* pv. *tomato* DC3000 genome should not be inferred from the numbers assigned.

[z] Locations of "*hop* islands" in *P. syringae* pv. *tomato* DC3000 (Wei et al. 2007), are indicated by dashed lines and are labeled with Roman numerals.

Table 2. *Continued from preceding page*

VR[x]	*P. syringae* pv. *syringae* B728a[y]		*P. syringae* pv. *tomato* DC3000		*P. syringae* pv. *phaseolicola* 1448A		*hop islands*[z]
	Gene	Locus	Gene	Locus	Gene	Locus	
	hopAA1-1	1183	hopAA1-1	1372	hopAA1-1	1263	VI
	hrpW	1184	hrpW	1373	hrpW	1264	
	shcM, hopM1	1185 to 1186	shcM, hopM1	1374 to 1375	shcM, hopM1	1265 to 1266	
	shcE, avrE1	1187 to 1188	shcE, avrE1	1376 to 1377	shcE, avrE1	1267 to 1268	
36	hrpH	1189	hrpH	1378	hrpH	1269	
	hrpRS, hrpA-E, hrpF-V	1190 to 1202	hrpRS, hrpA-E, hrpF-V	1379 to 1391	hrpRS, hrpA-E, hrpF-V	1270 to 1281	
37	hrpJ-hrcU operon	1205 to 1216	hrpJ-hrcU operon	1392 to 1403	hrpJ-hrcU operon	1282 to 1293	
	hrpL, hrpK	1217 to 1218	hrpL, hrpK	1404 to 1405	hrpL, hrpK	1294 to 1295	
	avrB3	1219					
	hopX1	1220			hopX1	1296	
	hopZ3	1224	hopB1	1406			
	Chaperone (put.)	1225	3 hypothet. genes	5622, 1408 to 1409			
38			Levansucrase	1453			
40	hopAF1	3813	hopAF1	1568	hopAF1	1443	
	MarR regulator	3737	MarR regulator	1645	MarR regulator	1519	
42	NRPS	3722					
44			TBDR	1855	hopF3	3498	
48	rulAB	1884 to 1885					
	Hemolysin (put.)	1886					
	hopH1	1889					
	hopAP1	1890					
	apbE family	1900	apbE family	2105	apbE family	1855	
49					TBDR	1870	
	Pyoverdin	1945 to 1962	Pyoverdin	2134 to 2153	Pyoverdin	1910 to 1929	
59					TBDR	2849	
60			T6SS	2539, 2546, 2549, 2553 to 2554			
62			Yersiniabactin	2595 to 2602	Yersiniabactin	2897 to 2904	
			TBDR	2605, 2607	TBDR	2892, 2894	
64					T3SS	2519 to 2521, 2524, 2527 to 2530, 2534 to 2538	
	Tc toxin-TccC	2412	hopP1	2678	Tc toxin-TccC	2571	
66	Phaseolotoxin (deg.)	2549 to 2555			Phaseolotoxin (deg.)	2705 to 2711	
	Syringofactins	2576 to 2577	Syringofactins	2829 to 2830	NRPS (deg.)	2729	
67-B					TBDR	2749	
					Achromobactin	2751 to 2760	
68-B	TBDR	2582	TBDR	2846			
	Achromobactin	2584 to 2593	TBDR	2853			
	Syringomycin/ syringopeptin	2608 to 2616					
	Streptomycin resist.	2669 to 2670					
70	rulAB	2761 to 2762					
					shcV-hopV1	2351 to 2352	
72			hopAB2	3087	hopAB3'	2294	
74					pel (deg.)	2184	
					T2SS	2173 to 2183	
75			FHA	3210			
			FHA (deg.)	3214			
			FHA (intein)	3229			
	Hemolysin	3089					
	aprA-like	3163	aprA-like	3332	aprA-like (deg.)	3076	
76					avrB4-1	3028	
	hopAH2	3123	hopAH2-1	3292	hopAH2	3036	VII
			hopAH2-2	3293			
	T2SS	3141 to 3151	T2SS	3307 to 3317	T2SS	3054 to 3064	
	Capsular polysacch.	3301 to 3308, 3231 to 3232	Capsular polysacch.	3222 to 3227, 3449 to 3450	Capsular polysacch.	3529 to 3534, 3150 to 3151	
	psl EPS	3301 to 3311	psl EPS	3529 to 3539	psl EPS	3222 to 3232	
80	plcA1 lipase	1823	plcA1 lipase	3648	plcA1 lipase	1783	
81	NRPS	1790 to 1796			NRPS	1749, 1750, 1751	
85	Syringolin biosynth.	1701 to 1706					
90-B			pgl	3960	pgl (deg.)	1530	
	RulAB	1424 to 1425			Shikimate kinase	1525	
	Cu resist.	1493 to 1498					
	Type IV pilus biogen.	1509 to 1518					
91-B					NRPS	3750	
					Oxidoreductase, ribD, hydrolase	3757 to 3759	

Continued on following page

Table 2. *Continued from preceding page*

VR[x]	*P. syringae* pv. *syringae* B728a[y]		*P. syringae* pv. *tomato* DC3000		*P. syringae* pv. *phaseolicola* 1448A		*hop* islands[z]
	Gene	Locus	Gene	Locus	Gene	Locus	
93			**avrPto1**	**4001**			
	hopAK1	**3839**	**hopAK1**	**4101**	**hopAK1**	**1424**	
100	pnlA	3985	pnlA	4283	pnlA	3992	
	Binary cyotoxin	3989 to 3990	Binary cytotoxin	5625, 4287	Binary cytotoxin	3996 to 3997	
101			**hopE1**	**4331**			
102	Tc toxin cluster	4032 to 4038	Tc toxin cluster	4341 to 4344	Tc toxin cluster (deg.)	4042 to 4045	
104	Tc toxin-TcdA1	4205	NRPS	4518 to 4519			
105	Binary cytotoxin	4244 to 4245	Binary cytotoxin	4570 to 4571	Binary cytotoxin	4269, 4271	
	opuC transport	4249 to 4252	opuC transport	4575 to 4578	opuC transport	4276 to 4279	
106			**hopS2, shcS2**	**4588** to **4589**			VIII
			hopT2	**4590**			
			hopO1-3'	**4592**			
			hopT1-2	**4593**			
			hopO1-2	**4594**			
			hopS1', shcS1	**4597, 4599**			
					argK-phaseolotoxin	4301 to 4319	
	hopAE1	**4269**			**hopAE1**	**4326**	
107			cfa	4680 to 4690			
			hopAD1	**4691**			
			NRPS	4699			
			hopAQ1	**4703**			
			cma	4704 to 4714			
			hopAA1-2	**4718**			IX
			hopV1, shcV	**4720** to **4721**			
			hopAO1	**4722**			
			hopD'	**4724**			
			hopG1	**4727**			
			hopH'	**5623**			
			hopQ1-2	**4732**			
			hypothet.	**4733**			
	hopI1	**4326**	**hopI1**	**4776**	**hopI1**	**4366**	
	pilGHIJ	488 to 491	pilGHIJ	5031 to 5034	pilGHIJ	0479 to 0482	
	hopAN1	**0465**	**hopAN1**	**5061**	**hopAN1**	**0456**	
	Hemolysin III	0452	Hemolysin III	5077	Hemolysin III	0443	
	pilMNOPQ	0407 to 0411	pilMNOPQ	5128 to 5132	pilMNOPQ	0390 to 0394	
111	Tat pathway	0382 to 0384	Tat pathway	5155 to 5157	Tat pathway	0365 to 0376	
115			**shcA, hopA1**	**5353** to **5354**			
116	**avrPto1**	**4919**					
117					**Glycosyl hydrolase**	**4991**	
					Levansucrase	4994	
118			T6SS	5418, 5419, 5424, 5427, 5435			
	Mangotoxin biosyn.	5005 to 5011	Mangotoxin biosyn.	5452 to 5458	Mangotoxin biosynth.	5084 to 5090	
120	Catalase katN	5095			Catalase katN	5178	
pDC3000A			**hopAM1-2**	**A0005**			X
			hopX1	**A0012**			
			hopO1-1	**A0018**			
			hopT1-1	**A0019**			
			Levansucrase	A0032			
pDC3000B			plcA2 lipase	B0005			
P1448A-A					pgl	A0006	
					hopW1-1	**A0009**	
					hopD1	**A0010**	
					Hypothet.	**A0011**	
					hopQ1	**A0012**	
					pgl	A0022	
					Levansucrase	A0027	
					hopAU1	**A0031**	
					hopAV1	**A0056**	
					saf fimbriae	A0062 to 0065	
					hopW1-2	**A0075**	
					avrRps4	**A0087**	
					avrD1	**A0113**	
					avrB2	**A0120**	
					hopAW1	**A0122**	
					hopAB1	**A0127**	
					Shikimate kinase	**A0133**	
					Hypothet.	**A0134**	
p1448A-B					rulAB	B0003 to 0004	

in the genomes in which they occur. Though the syntenic conservation of their flanking regions argues strongly for ancient association with the *P. syringae* lineage, several of these *hop* genes appear to have undergone more recent rearrangement. The *hopAG1-AH1-AI1* cluster in VR $24_{Pto\ DC3000}$ is disrupted by an IS*Pssy* insertion in *hopAG1*, *hopAF1*$_{Pto\ DC3000}$ in VR 40 is inverted relative to its orthologs in *P. syringae* pv. *phaseolicola* 1448A and *P. syringae* pv. *syringae* B728a, and a duplication event appears to have occurred in *P. syringae* pv. *tomato* DC3000 resulting in the presence of a *hopAH2* paralog immediately upstream of the syntenically conserved *hopAH2-2*$_{Pto\ DC3000}$ in VR 76. The more ancient associations of *avrE*, *hopI1*, *AA1*, *AJ1*, and *AK1* implied by gene synteny are consistent with the results of Rohmer and associates (2004) based on analysis of codon usage and GC content (Rohmer et al. 2004).

Among *hop* genes shared by two or more strains but lacking syntenic conservation are *hopX1*, found in the exchangeable effector locus (VR 37) of *P. syringae* pv. *phaseolicola* 1448A and *P. syringae* pv. *syringae* B728a, and on *P. syringae* pv. *tomato* DC3000 plasmid pDC3000A. Members of the *hopAB* family are also found in nonsyntenic locations, with *hopAB1*$_{Pph\ 1448A}$ present on 1448A plasmid p1448A-A, *hopAB1*$_{Psy\ B728a}$ in VR 15, and the more closely related *hopAB2*$_{Pto\ DC3000}$ and *AB3*$_{Pph\ 1448A}$ in VR 72. *HopR1*, *G1*, *H1*, *V1*, *D1*, and *Q1* are limited to two of the three genomes, with orthologs, excepting *hopV1*$_{Pph\ 1448A}$, present in diverse VR (Table 2). Nonsyntenic distribution patterns are consistent with either independent acquisition by one or more strains or ancient acquisition followed by more recent intragenomic rearrangement.

While linkages observed for genes in the *avrE-hopM1-hopAA1*, *hopAJ-hopAT*, and *hopAG-AH-AI* clusters are most likely due to vertical inheritance of ancient acquisitions, linkage patterns can also be observed for selected clusters of nonsytenically conserved *hops* and members of the Hrp-regulon, consistent with comobilization during either horizontal transfer or intragenomic duplication. Examples include the HrpL-regulated operon PSPTO_0835 to 0837, which, together with 575 bp of conserved upstream DNA containing the HrpL promoter, is found at VR $24_{Pto\ DC3000}$ and $91_{Pph\ 1448A}$, and members of the *hopO* and *hopT* families, found duplicated in VR $106_{Pto\ DC3000}$ and on pDC3000A.

Comobilization also appears likely for *hopD1* and *hopQ1-1*, located in VR $24_{Pto\ DC3000}$ and on p1448A-A and conserved together with 340 and 90 bp of flanking DNA containing their HrpL promoters. The relatively close proximity of *hopD* and *hopQ* orthologs in VR $107_{Pto\ DC3000}$ raises the possibility that the two genes were comobilized to that region as well, with subsequent insertions of *hopG1*, *hopH1*, and two insertion sequences leading to the 6.4 kb presently separating the two genes. Interestingly, BLASTP analysis of *hopD1* ortholog sequences deposited in GenBank reveals the presence of flanking sequences similar to *hopQ1* in both *P. syringae phaseolicola* 1302A (AJ277494) and *Pantoea agglomerans* pv. *gypsophila* (AF462346). The persistence of this linkage in diverse genera suggests that the biological roles and possible functional association of these two proteins may be influencing a broad range of plant-bacterial interactions. Though deletion of *hopQ1-1* has been shown to expand host range for *P. syringae* pv. *tomato* DC3000 (Wei et al. 2007), neither the role of HopD1 nor the possibility of a HopD1-HopQ1 functional interaction has been explored.

In contrast to the members of the HrpL regulon, which strongly segregate with VR and are not generally represented across all three genomes (T3SS structural genes excepted), genes exhibiting reduced activity in the absence of transcription activators HrpRS, independent of HrpL-dependent activation, generally segregate with the conserved core of the *P. sy-*

ringae genomes. Of the 60 genes thus identified (Lan et al. 2006), only 24 are present in VR, and all but 16 are syntenically conserved among the three genomes.

Toxins exhibit divergent distribution patterns depending on association with plant and insect pathogenesis. The most striking feature of toxin gene distribution is the divergent distribution patterns observed between those known only from plant pathogens and those with insect-active orthologs. Confirmed phytotoxins coronatine (VR $107_{Pto\ DC3000}$), syringomycin (VR $107_{Psy\ B728a}$), and phaseolotoxin (VR $106_{Pph\ 1448A}$) are strain-specific in their localization, and the degenerate phaseolotoxin cluster is shared by only two of the three strains (VR 66).

In contrast, toxin genes shared with insect pathogens—the two *xaxAB* loci (Vigneux et al. 2007), the mangotoxin cluster (Arrebola et al. 2007), and the toxin complex c cluster (Waterfield et al. 2001)—even when associated with VR, are syntenically conserved across all three genomes, suggestive of ancient association with the *P. syringae* lineage. A similar pattern is observed for the AprA metalloprotease and chitinase. Future investigation of this growing body of insect-associated virulence factors may reveal new perspectives on *P. syringae* niche adaptation. Does their presence indicate a long-standing and potentially ongoing ability to occupy an insect-associated niche as was recently demonstrated for the soft-rot phytopathogen *Dickeya dadantii* (Grenier et al. 2006)? Or do these virulence factors, like mangotoxin, have a broad spectrum of activity that *P. syringae* employs primarily in its interaction with plants?

Siderophores, adhesins, secretion pathways, and other virulence factors—segregation between core and variable genome points to possible sources of strain-specific interaction with host plants. The complex locus encoding genes for the biosynthesis and transport of the siderophore pyoverdine is syntenically conserved in the core *P. syringae* genome, consistent with its widespread presence in diverse pseudomonads. In contrast, achromobactin (*P. syringae* pv. *phaseolicola* 1448A and *P. syringae* pv. *syringae* B728a) and yersiniabactin (*P. syringae* pv. *tomato* DC3000 and *P. syringae* pv. *phaseolicola* 1448) loci are variable in their distribution, though the complexity of the regions in which they occur and their location near boundaries of genome rearrangements makes definitive judgment of syntenic conservation difficult. Genes encoding predicted TBDR (Blanvillain et al. 2007) are found closely linked to biosynthetic clusters for the three siderophore types. There appear to be over 15 TBDR syntenically conserved among the three genomes, with five and four additional unique TBDR present in VR of the DC3000 and 1448A genomes, respectively. Roles for the strain-specific TBDR have not been explored, but they could impact fitness at various stages of the bacterial-plant interaction through uptake of siderophores and other substrates.

In addition to those with characterized roles in toxin and siderophore production, seven other predicted nonribosomal peptide synthases are found in the *P. syringae* genomes. All are found in VR, with five conserved in neither location nor sequence, suggestive of possible strain-specific roles.

Comparison of factors with confirmed and potential involvement in adhesion also reveals significant differences among the three strains. All three contain syntenically conserved clusters encoding genes for biogenesis of alginate, Psl exopolysaccharides, type IV pili and, though apparently degenerate in *P. syringae* pv. *phaseolicola* 1448A, syringofactins. However, *P. syringae* pv. *tomato* DC3000 and 1448A encode additional complements of adhesin-related genes, segregating with the VR. In DC3000, these include the *wss* cellulose biosynthesis cluster (VR 27) and genes encoding two nondegenerate filamentous hemagglutinins (FHA) (VR 75), and, in 1448A, a

plasmid-borne cluster of genes associated with biogenesis of *saf* atypical fimbriae and genes encoding a hemagglutinin (VR 24) and FHA (VR 19) distantly orthologous to those found in DC3000. The divergent repertoires of predicted adhesins may play determining roles in the relative attachment capacity observed for the three strains.

Among the more recently identified secretion pathways, the Tat pathway cluster, like the T3SS, is located in a region exhibiting compositional deviation but conserved in all three genomes (VR 111), as is the chromosomally borne gene for its substrate PlcA1 lipase. In contrast, the two T6SS (Mougous et al. 2006; Pukatzki et al. 2006) clusters are limited to DC3000 and are found in VR 60 and 118, suggesting a strain-specific role if shown to be functional.

While the majority of genes implicated in osmotic and oxidative stress tolerance are syntenically conserved, some differences are evident. *P. syringae* pv. *phaseolicola* 1448A and *P. syringae* pv. *syringae* B728a share two catalases (in VR 19 and 120) not found in *P. syringae* pv. *tomato* DC3000, and B728a encodes four orthologs of the *rulAB* DNA polymerase involved in repair of UV damage (VR 24, 48, 70, 90) not present in either DC3000 or 1448A. The additional copies of the *rulAB* locus have been previously implicated in the relative fitness of B728a on the epiphytic surface (Feil et al. 2005).

The collective distribution patterns for virulence and host-association genes support a model in which a commonly held foundation of syntenically conserved genes predisposes diverse strains for exploitation of the plant niche, with additional strain-specific repertoires, localized to the VR, shaping distinctive aspects of each strain's interaction with its plant hosts. That said, experimental confirmation of expression and functions are critical to the ultimate role assignments for individual gene products. As illustrated by the recent work with yersiniabactin, functional characterization can prove surprisingly difficult, even when bioinformatics has provided compelling hypotheses (Jones et al. 2007).

The composition of VR is shaped by diverse mobile genetic elements.

Content of VR is typically shaped by the combined effects of rearrangement, disruption, deletion, and horizontal transfer (Gal-Mor and Finlay 2006; Hacker and Kaper 2000). Examination of the syntenically divergent VR reveals enrichment for mobile genetic elements, with 82% of the chromosomally-borne DC3000 mobile genetic elements located in VR. Among these are integrative and conjugative elements, prophages, phage integrases, and insertion elements (IS), analysis of which provides clues to the mechanisms underlying the composition of specific VR.

For example, the content of VR 24, $90_{Psy\ B728a}$, and $91_{Pph\ 1448A}$ appears to have been largely shaped by integrative and conjugative elements (ICElands) (Burrus et al. 2002). PPHGI-1, the first ICEland characterized in *P. syringae*, was shown by Jackson and associates (2000) to excise from the genome of *P. syringae* pv. *phaseolicola* 1302A, resulting in deletion of the effector gene *hopAR1* and expansion of the 1302A host range (Jackson et al. 2000; Pitman et al. 2005). Sequences orthologous to PPHGI-1 were also reported for *P. syringae* pv. *tomato* DC3000 and *P. syringae* pv. *syringae* B728a genomes (Pitman et al. 2005). As shown in Figure 1A, close comparison of PPHGI-1 with the sequenced *P. syringae* genomes reveals elements of the conserved ICEland framework in VR 24 for all three genomes, interspersed with variable content that includes numerous virulence-associated genes. VR $24_{Pto\ DC3000}$ contains the most extended regions of similarity to previously characterized ICElands, interspersed with *hopD1*, *Q1-1*, *AJ1*, *AT1*, *R1* genes, and the HrpL-regulated operon PSPTO_0834 to 0837. VR $24_{Pph\ 1448A}$ carries the *hopAJ1*, *hopAT1*, *hopG1*, *avrB4-2* genes as well as genes for two type IV pilins, two fimbrial proteins, and a hemagglutinin. Though the only conserved element of the ICEland framework in VR $24_{Psy\ B728a}$ is a phage integrase syntenically conserved with those in the other genomes, it is closely linked to *avrRpm1* and one of the extra copies of *rulAB* previously described for B728a (Feil et al. 2005) (Fig. 1A).

VR $90_{Psy\ B728a}$ also includes regions of extended similarity with PPHGI-1 interspersed with genes predicted to confer copper resistance, a gene cluster for type IV pili biogenesis, and yet another of the extra *rulAB* loci. A phage integrase orthologous to that at the ICEland flank is present in VR $91_{Pph\ 1448A}$ and is associated with HrpL-regulated operon PSPTO_ 3759 to 3756, though no other conserved ICEland elements are present. The insertions at VR 24, 90, and 91 are associated with tRNAs, frequently cited as hotspots for insertion of phage integrases.

Distribution of ICEland sequences among diverse genomes was further explored by BLASTP comparison of DC3000 coding sequences against the predicted gene products of 58 other bacterial genomes, including 12 sequenced pseudomonads, representatives of all sequenced genera in the γ and β-proteobacteria, and including completed sequences of the "canonical" phytopathogens *Pectobacterium atrosepticum*, *Xyllela fastidiosa*, *Xanthomonas campestris* pv. *vesicatoria*, *Xanthomonas campestris* pv. *campestris* ATCC 33913, *Xanthomonas axonopodis* pv. *citri*, *Xanthomonas oryzae* pv. *oryzae*, *Burkholderia cenocepacia* HI2424, and *Ralstonia solanacearum*. As shown in Figure 1B, BLASTP analysis reveals extended similarity with ICEland-associated genes, not only in *P. syringae* pv. *syringae* B728a but also in *P. aeruginosa* PA14, *Ralstonia metallidurans* CH34, *Burkholderia xenovorans* LB400, *Salmonella enterica* subsp. *enterica* serovar Paratyphi A ATCC 9150, *Photorhabdus luminescens* subsp. *laumondii* TTO1, and phytopathogens *Xanthomonas campestris* pv. *vesicatoria* 85-10, and *P. atrosepticum* SCRI1043. Interestingly, the well-conserved ICEland found in *P. atrosepticum* encodes genes orthologous to the *cfa* operon required for coronafacic acid biosynthesis in *P. syringae* pv. *tomato* DC3000 (Fig. 1A). The ICEland region identified in *Xanthomonas campestris* pv. *vesicatoria* 85-10 spans loci XCV2360-2414, with variable content corresponding to loci XCV2382 to XCV2385 and XCV2402 to XCV2406. None of the genes in this region have been implicated in virulence or host association.

Though several prophage clusters are found among the three *P. syringae* genomes, *avrPto1* appears to be the only prophage-associated virulence gene. $AvrPto1_{Pto\ DC3000}$ is located immediately downstream of a nearly complete prophage cluster (VR 93), $avrPto1_{Psy\ B728}$ is linked to orthologous prophage components at a nonsyntenic location (VR 116), and comparison of $avrPto1_{Pto\ DC3000}$ with the GenBank deposition for $avrPto1_{Pto\ JL1065}$ reveals that orthologous phage genes are also present immediately adjacent to $avrPto1_{Pto\ JL1065}$. Collectively, these observations suggest that prophage-mediated mobilization is the dominant mechanism for intergenome mobilization of *avrPto1*.

hopH1 and *hopC1* in VR $18_{Pto\ DC3000}$ are linked to a putative prophage-derived pyocin insert between *trpE* and *trpG*. Although this association is not found elsewhere in the three sequenced strains, analysis of GenBank deposition for the *hopC* orthologs in *P. syringae* pv. *pisi* race 1 (299A) (AJ277496) reveals a conserved arrangement between orthologs of *hopH1*, *hopC1*, and two genes immediately flanking. Whether this linkage results from vertical inheritance of a syntenically conserved region or pyocin-associated mobilization to diverse locations is presently unknown.

Insertion sequences acting in concert as composite transposons have also played a significant role in shaping the con-
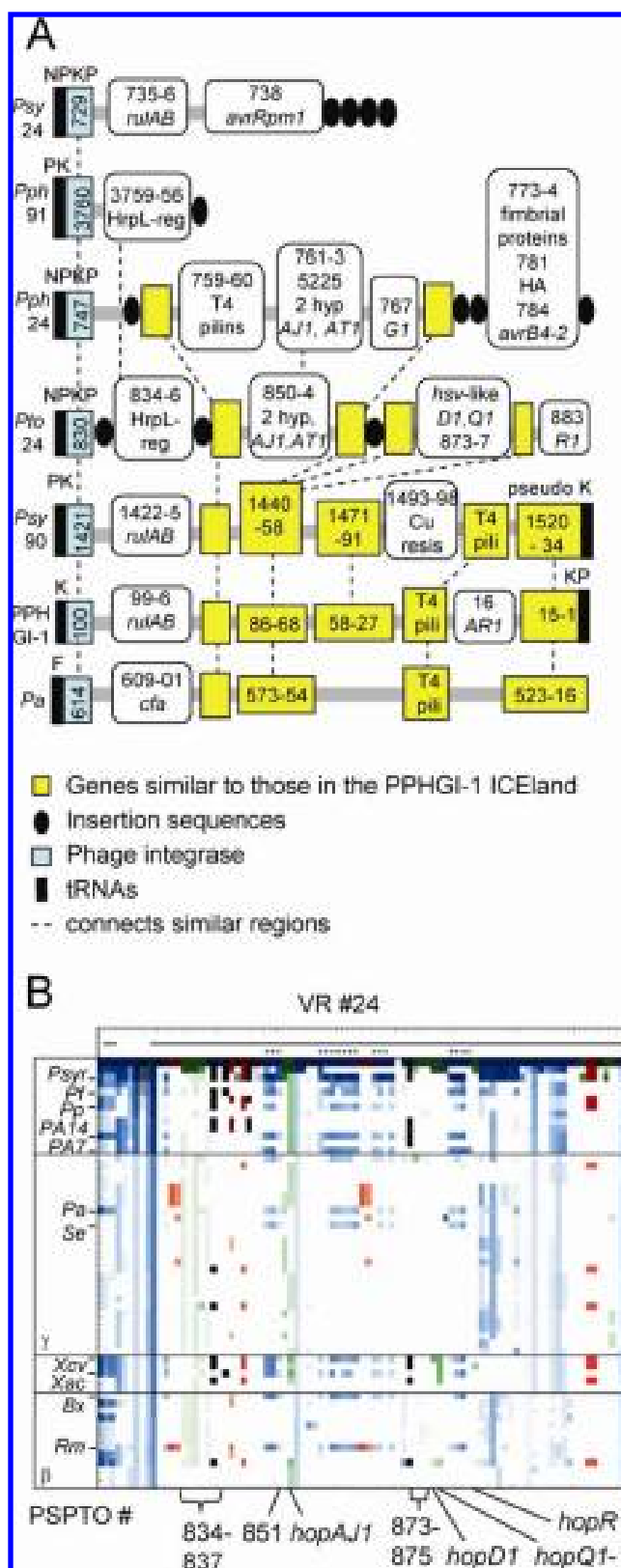
tent of VR. Ongoing analyses are revealing large numbers of gene fragments within the VR, linked in many instances to the locations of IS insertions. An example of this phenomenon is readily observable in VR 107$_{Pto\ DC3000}$, in which a *hopD* ortholog has been disrupted by an IS*52* insertion and a *hopH* ortholog by an IS*Psy4* insertion embedded in an insertion of IS*Pssy*. Gene fragments resulting from IS insertions may hold the key to deciphering locations of discrete insertion events that have shaped VR content, as illustrated by *P. syringae* pv. *tomato* DC3000 VR 1, in which locations of the N and C-terminal fragments of an Mg-chelatase gene (PSPTO_0042 and PSPTO_0060) indicate that PSPTO_0043 to 0059 (which includes *hopK1*$_{Pto\ DC3000}$) were transferred to this region of the genome during a single mobilization event.

VR are enriched for genes of low and atypical distribution relative to other bacterial genomes.

Given the association of VR with confirmed virulence factors of nonuniform distribution and mobile elements promoting gene transfer and rearrangement, these regions are likely enriched for additional, as-yet-uncharacterized genes contributing to the bacterium-host interaction. To evaluate the distribution of predicted gene products in the *P. syringae* pv. *tomato* DC3000 VR among bacteria occupying diverse niches and of varying degrees of genetic relatedness, chromosomal and plasmid-borne coding sequences from DC3000 were compared with coding sequences of 41 other bacterial genomes, including 12 sequenced pseudomonads, representatives of all sequenced genera in the γ and β-proteobacteria, and including completed sequences of the canonical phytopathogens. Results of BLASTP analyses are shown in the wheel diagrams in Figure 2, generated using Genome Diagram software (Pritchard et al. 2006). Locations of DC3000 VR are indicated in the outermost ring of the genome wheel. Similar representations of large-scale genome comparison have previously been diagrammed for *P. atrosepticum* (Toth et al. 2006). A linear representation of BLASTP data for the DC3000 chromosome and plasmids, similar to the output shown in Figure 1B is provided on the Pseudomonas-Plant Interaction website.

Consistent with observations made for *P. atrosepticum* (Bell et al. 2004; Toth et al. 2006), the *P. syringae* pv. *tomato* DC3000 genome is a mosaic of broadly conserved regions interspersed with relatively open "rays" reflecting low or inconsistent conservation of component genes across the majority of genomes analyzed. VR generally correlate with locations of these open rays, though some also encompass areas of greater cross-genome conservation, particularly when component genes are members
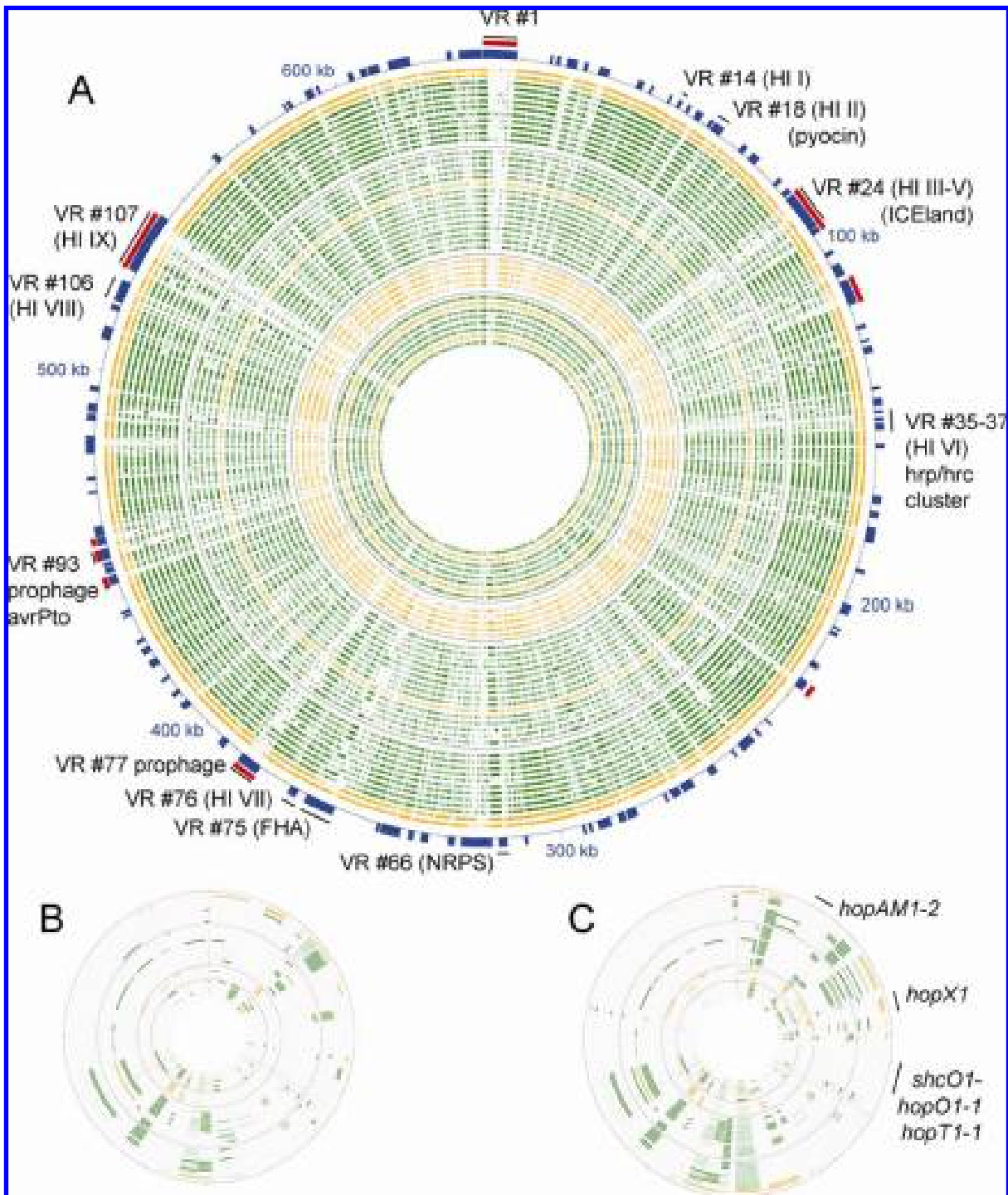

Fig. 1. Conservation between *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*) variable region 24 (VR #24) and regions in other genomes. **A,** Comparison of DC3000 VR 24 with the integrative and conjugative element (ICEland) PPHGI-1 from *P. syringae* pv. *phaseolicola* 1302A and specified regions in *P. syringae* pv. *phaseolicola* (*Pph*), *P. syringae* pv. *syringae* (*Psy*), and *Pectobacterium atrosepticum* (*Pa*). cfa = coronafacic acid, Cu resis. = copper resistance, HA = hemagglutinin, hrpL-reg. = genes in the hrpL regulon, hyp. = hypothetical, and T4 pili = type 4 pili. Hops AJ, AT, AR1, D1, G1, Q1, and R1 are shown without the *hop* gene designation. **B,** Presence of DC3000 VR 24 orthologs in 58 other genera as determined by BLASTP analysis. Orthology of DC3000 loci (shown in sequential order on the *x* axis) with open reading frames in selected bacterial genomes (listed on the *y* axis) is indicated with colored squares, with the intensity of color proportional to the degree of sequence similarity. Mobile genetic elements are shown in red and members of the HrpL regulon in green. Pseudogenes, evaluated by BLASTN due to the absence of coding sequences in the genome annotation, are shown in black. Loci corresponding to ICEland components conserved between VR 24$_{Pto\ DC3000}$ and PPHGI-1 are indicated with asterisks. Genomes are grouped into four categories corresponding to pseudomonads, nonpseudomonal γ-proteobacteria, xanthomonads, and β-proteobacteria. Individual genomes showing orthology to the ICEland sequences are indicated as follows: *Psyr* = *P. syringae* pv. *syringae* B728a, *Pf* = *P. fluorescens* Pf-5, *Pp* = *P. putida* F1, *PA14* = *P. aeruginosa* UCBPP-PA14, *PA7* = *P. aeruginosa* PA7, *P Pa* = *ectobacterium atrosepticum* SCRI1043, *Se* = *Salmonella enterica* subsp. *enterica* serovar Typhi CT18, *Xcv* = *X. campestris* pv. *vesicatoria* 85-10, *Xac* = *X. axonopodis* pv. *citri* 306, *Bx* = *Burkholderia xenovorans* LB400, *Rm* = *Ralstonia metallidurans* CH34. A complete listing of the 58 genomes screened is provided in the supplemental information published online.

Fig. 2. Representation of *Pseudomonas syringae* pv. *tomato* DC3000 coding sequences in selected bacterial genomes as determined by BLASTP analysis of **A**, the DC3000 chromosome (NC_004578), **B**, DC3000 plasmid pDC3000B (NC_004632), and **C**, DC3000 plasmid pDC3000A (NC_004633). **A**, The outermost ring represents DC3000, with purple bars on the chromosome wheel indicating locations of variable regions. Presence of orthologs in other genomes is indicated on the remaining rings, with intensity of color proportional to the degree of sequence similarity. Locations of the DC3000 *hop* islands are shown (Wei et al. 2007), and regions enriched for genes found in two or fewer of the genomes screened are indicated with red bars. Genomes are grouped into four categories corresponding to pseudomonads, nonpseudomonad γ-proteobacteria, xanthomonads, and β-proteobacteria, with each category delineated by a solid gray line. The complete list of genomes from outer to innermost ring are as follows. Pseudomonads: 1, *P. syringae* pv. *phaseolicola* 1448A; 2, *P. syringae* pv. *syringae* B728a; 3, *P. fluorescens* PfO-1; 4, *P. fluorescens* Pf-5; 5, *P. putida* KT2440; 6, *P. putida* F1; 7, *P. entomophila* L48; 8, *P. stutzeri* A1501; 9, *P. mendocina* YMP; 10, *P. aeruginosa* UCBPP-PA14; 11,, *P. aeruginosa* PAO1; 12, *P. aeruginosa* PA7. Nonpseudomonad γ–proteobacteria: 13, *Photorhabdus luminescens* subsp. *laumondii* TTO1, 14. *Shewanella frigidimarina*, 15. *Shewanella* sp. MR-4, 16. *Yersinia pseudotuberculosis* IP 32953, 17. *Yersinia pestis* Antiqua; 18, *Pectobacterium atrosepticum* SCRI1043; 19, *Salmonella enterica* subsp. *enterica* serovar Choleraesuis SC-B67; 20, *Salmonella enterica* subsp. *enterica* serovar Typhi CT18; 21, *Salmonella enterica* subsp. *enterica* serovar Paratyphi A ATCC 9150; 22, *Escherichia coli* O157:H7 EDL933; 23, *Escherichia coli* UTI89; 24, *Escherichia coli* K12; 25, *Vibrio vulnificus* YJ016; 26, *Vibrio cholerae* O1 biovar eltor N16961; 27, *Shigella flexneri* 5 8401; 28, *Aeromonas hydrophila* subsp. *hydrophila* ATCC 7966. Xanthomonads: 29, *X. campestris* pv. *vesicatoria* 85-10; 30, *X. campestris* pv. *campestris* ATCC 33913; 31, *X. axonopodis* pv. *citri* 306; 32, *X. oryzae* pv. *oryzae* MAFF 311018; 33, *Xylella fastidiosa* Temecula1. β-proteobacteria; 34, *Burkholderia xenovorans* LB400; 35, *Burkholderia cenocepacia* HI2424; 36, *Burkholderia thailandensis* E264; 37, *Burkholderia mallei* ATCC 23344; 38, *Ralstonia metallidurans* CH34; 39, *Ralstonia eutropha* JMP134; 40, *Ralstonia solanacearum* GMI1000; 41, *Bordetella bronchiseptica* RB50. Canonical phytopathogens are shown in orange.

of broadly conserved families. Examples include the genes encoding nonribosomal peptide synthase and filamentous hemagglutinin associated with VR 66 and 75, respectively, that elicit strong BLASTP hits from family members in many other species (Fig. 2A). As with the VR on the chromosomes, the plasmid coding sequences also exhibit low and idiosyncratic conservation with the contents of other genomes. Altogether, there are 832 DC3000 coding sequences with orthologs in no more than two of the 41 genomes analyzed. Of these, 438 are unique to DC3000 and 135 are limited to the three sequenced *P. syringae* genomes, representing not only candidate genes for strain and species-specific host interaction factors but also a valuable population of potential molecular markers for use in genome-based diagnostics (Vieira et al. 2007). Ongoing genome sequencing of additional *P. syringae* strains and other pseudomonads is expected to further refine the pools of strain- and species-specific genes. VR particularly enriched for genes of limited distribution include VR 1, 24, and 107 (Fig. 2A).

Aside from the virulence genes and mobile elements already noted, what functional information is available for genes in the VR? For the most part, very little is known, with over 30% of the genes in VR presently annotated as "hypothetical" or "conserved hypothetical." While the present review focuses on distribution patterns observed for known virulence-associated genes, future work will involve analysis of expression and distribution patterns for this uncharacterized reservoir of potential host-association genes. An important first step is represented by the ongoing direct sequencing of the *P. syringae* pv. *tomato* DC3000 transcriptome. This information on DC3000 transcripts is not only providing confirmation for gene products identified by orthology, such as the XaxAB toxins, but is also confirming expression and significantly refining coordinates for those genes lacking functional annotation (M. Filiatrault, *personal communication*). Complementary data obtained by shotgun proteomics methods confirm the expression of proteins encoded by what have, until now, been known only as hypothetical genes based on machine-generated ORF calls (D. J. Schneider, *personal communication*). These data, to be added to the annotation records at NCBI, will provide a sound experimental basis for accurate annotation of orthologs in strains and species with genome sequencing planned or underway and pave the way for meaningful investigation of their distribution among diverse bacteria and potential roles in the host-pathogen interaction.

Summary.

Advances in *P. syringae* gene characterization made since deposition of the original genome sequences for DC3000, 1448a, and B728a are broadening the current understanding of factors implicated in virulence and survival in the plant-associated niche.

To explore the relationship between virulence gene location and genome structure, a systematic approach for genome structure characterization was developed, incorporating compositional deviation and divergence in genome alignment. This approach, used here to generate a map of core regions and VR for the *P. syringae* genomes, is readily applicable to other sets of related genomes.

Analysis of the locations of virulence factors relative to the *P. syringae* genome map reveals strong segregation of the HrpL regulon with VR, divergent distribution patterns for toxin genes depending on their association with plant or insect pathogenesis, and patterns of distribution for other virulence genes that highlight potential sources of strain-to-strain difference in host interaction.

A subset of the VR contain dense mosaics of mobile genetic elements including phage, insertion sequences, and ICElands

interspersed with gene fragments resulting from their activity. Ongoing identification of gene fragments coupled with future comparisons with additional *P. syringae* genomes will allow researchers to more clearly distinguish the individual genome rearrangement events that have shaped genome structure.

BLASTP analysis of *P. syringae* pv. *tomato* DC3000 genes with 41 other sequenced genomes reveals that the VR are enriched for genes with limited distribution and little or no functional characterization. Their presence in regions of high gene turnover and horizontal transfer suggests a possible role in unexplored strategies of adaptation to host and other environmental niches. Further characterization will be accomplished by comparison with the full range of genes and functional motifs available in public databases, together with application of proteomics data and transcriptome mapping to verify expression and refine annotated boundaries.

The resulting high-value annotation record for DC3000 will represent an invaluable resource for in-depth experimental characterization of the biological roles of expressed but functionally uncharacterized genes in the VR and comparison of gene repertoire and syntenic arrangement with those of *P. syringae* genomes currently being sequenced.

Since publication of its genome sequence in 2003, the significance of *P. syringae* pv. *tomato* DC3000 as a model system has been repeatedly reinforced by extensive characterization of virulence determinants, large-scale bioinformatic analyses, and structural characterization of gene products by researchers in diverse fields. Ongoing curation of its genome annotation through incorporation of new data such as those reviewed here is vital to the continuing role of DC3000 as a model system for experimental investigation and a model annotation record upon which future *P. syringae* genome annotations can be based.

LITERATURE CITED

Abramovitch, R. B. and Martin, G. B. 2005. AvrPtoB: A bacterial type III effector that both elicits and suppresses programmed cell death associated with plant immunity. FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Lett. 245:1-8.

Alfano, J. R., Charkowski, A. O., Deng, W.-L., Badel, J. L., Petnicki-Ocwieja, T., van Dijk, K., and Collmer, A. 2000. The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. Proc. Natl. Acad. Sci. U.S.A. 97:4856-4861.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

Arrebola, E., Cazorla, F. M., Duran, V. E., Rivera, E., Olea, F., Codina, J. C., Perez-Garcia, A., and de Vicente, A. 2003. Mangotoxin: A novel antimetabolite toxin produced by *Pseudomonas syringae* inhibiting ornithine/arginine biosynthesis. Physiol. Mol. Plant Pathol. 63:117-127.

Arrebola, E., Cazorla, F. M., Romero, D., Perez-Garcia, A., and de Vicente, A. 2007. A nonribosomal peptide synthetase gene (*mgoA*) of *Pseudomonas syringae* pv. *syringae* is involved in mangotoxin biosynthesis and is required for full virulence. Mol. Plant-Microbe Interact. 20:500-509.

Beck, B. J., and Downs, D. M. 1998. The *apbE* gene encodes a lipoprotein involved in thiamine synthesis in *Salmonella typhimurium*. J. Bacteriol. 180:885-891.

Badel, J. L., Shimizu, R., Oh, H.-S., and Collmer, A. 2006. A *Pseudomonas syringae* pv. *tomato* avrE1/hopM1 mutant is severely reduced in

growth and lesion formation in tomato. Mol. Plant-Microbe Interact. 19:99-111.

Bell, K. S., Sebaihia, M., Pritchard, L., Holden, M. T., Hyman, L. J., Holeva, M. C., Thomson, N. R., Bentley, S. D., Churcher, L. J., Mungall, K., Atkin, R., Bason, N., Brooks, K., Chillingworth, T., Clark, K., Doggett, J., Fraser, A., Hance, Z., Hauser, H., Jagels, K., Moule, S., Norbertczak, H., Ormond, D., Price, C., Quail, M. A., Sanders, M., Walker, D., Whitehead, S., Salmond, G. P., Birch, P. R., Parkhill, J., and Toth, I. K. 2004. Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110.

Bender, C. L., Alarcon-Chaidez, F., and Gross, D. C. 1999. *Pseudomonas syringae* phytotoxins: Mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. Microbiol. Mol. Biol. Rev. 63:266-292.

Berti, A. D., Greve, N. J., Christensen, Q. H., and Thomas, M. G. 2007. Identification of a biosynthetic gene cluster and the six associated lipopeptides involved in swarming motility of *Pseudomonas syringae* pv. *tomato* DC3000. J. Bacteriol. 189:6312-6323.

Blanvillain, S., Meyer, D., Lauber, E., and Arlat, M. 2007. Plant carbohydrate scavenging through TonB-dependent receptors: A feature shared by phytopathogenic and aquatic bacteria. PLoS ONE 2:e224. Published online.

Boncompagni, E., Osteras, M., Poggi, M.-C., and le Rudulier, D. 1999. Occurrence of choline and glycine betaine uptake and metabolism in the family rhizobiaceae and their roles in osmoprotection. Appl. Environ. Microbiol. 65:2072-2077.

Bretz, J. R., Mock, N. M., Charity, J. C., Zeyad, S., Baker, C. J., and Hutcheson, S. W. 2003. A translocated protein tyrosine phosphatase of *Pseudomonas syringae* pv. *tomato* DC3000 modulates plant defence response to infection. Mol. Microbiol. 49:389-400.

Bronstein, P. A., Marrichi, M., Cartinhour, S., Schneider, D. J., and DeLisa, M. P. 2005. Identification of a twin-arginine translocation system in *Pseudomonas syringae* pv. *tomato* DC3000 and its contribution to pathogenicity and fitness. J. Bacteriol. 187:8450-8461.

Buell, C. R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I. T., Gwinn, M. L., Dodson, R. J., Deboy, R. T., Durkin, A. S., Kolonay, J. F., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M. J., Haft, D. H., Nelson, W. C., Davidsen, T., Zafar, N., Zhou, L., Liu, J., Yuan, Q., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Utterback, T., Van Aken, S. E., Feldblyum, T. V., D'Ascenzo, M., Deng, W. L., Ramos, A. R., Alfano, J. R., Cartinhour, S., Chatterjee, A. K., Delaney, T. P., Lazarowitz, S. G., Martin, G. B., Schneider, D. J., Tang, X., Bender, C. L., White, O., Fraser, C. M., and Collmer, A. 2003. The complete genome sequence of the Arabidopsis and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186.

Burrus, V., Pavlovic, G., Decaris, B. and Guedon, G. 2002. Conjugative transposons: The tip of the iceberg. Mol. Microbiol. 46:601-610.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M.-A., Barrell, B. G., and Parkhill, J. 2005. ACT: The Artemis comparison tool. Bioinformatics 21:3422-3423.

Chang, J. H., Urbach, J. M., Law, T. F., Arnold, L. W., Hu, A., Gombar, S., Grant, S. R., Ausubel, F. M., and Dangl, J. L. 2005. A high-throughput, near-saturating screen for type III effector genes from *Pseudomonas syringae*. Proc. Natl. Acad. Sci. U.S.A. 102:2549-2554.

Chatterjee, A., Cui, Y., and Chatterjee, A. K. 2002. Regulation of *Erwinia carotovora hrpL*Ecc sigma-LEcc, which encodes an extracytoplasmic function subfamily of sigma factor required for expression of the HRP regulon. Mol. Plant-Microbe Interact. 15:971-980.

Chatterjee, A., Cui, Y., Hasegawa, H., and Chatterjee, A. K. 2007. PsrA, the *Pseudomonas* sigma regulator, controls regulators of epiphytic fitness, quorum-sensing signals, and plant interactions in *Pseudomonas syringae* pv. *tomato* strain DC3000. Appl. Environ. Microbiol. 73:3684-3694.

Chang, W.-S., van de Mortel, M., Nielsen, L., Nino de Guzman, G., Li, X. and Halverson, L. J. 2007. Alginate production by *Pseudomonas putida* creates a hydrated microenvironment and contributes to biofilm architecture and stress tolerance under water-limiting conditions. J. Bacteriol. 189:8290-8299.

Chen, C., and Beattie, G. A. 2007. Characterization of the osmoprotectant transporter OpuC from *Pseudomonas syringae* and demonstration that CBS domains are required for its osmoregulatory function. J. Bacteriol. 189:6901-6912.

Chen, C., and Beattie, G. A. 2008. *Pseudomonas syringae* BetT is a low affinity choline transporter that is responsible for superior osmoprotection by choline over glycine betaine. J. Bacteriol. 190:2717-2715.

Donadio, S., Monciardini, P., and Sosio, M. 2007. Polyketide synthases and nonribosomal peptide synthetases: The emerging view from bacterial genomics. Nat. Prod. Rep. 24:1073-1109.

Espinosa, A., Guo, M., Tam, V. C., Fu, Z. Q., and Alfano, J. R. 2003. The *Pseudomonas syringae* type III-secreted protein HopPtoD2 possesses protein tyrosine phosphatase activity and suppresses programmed cell death in plants. Mol. Microbiol. 49:377-387.

Feil, H., Feil, W. S., Chain, P., Larimer, F., DiBartolo, G., Copeland, A., Lykidis, A., Trong, S., Nolan, M., Goltsman, E., Thiel, J., Malfatti, S., Loper, J. E., Lapidus, A., Detter, J. C., Land, M., Richardson, P. M., Kyrpides, N. C., Ivanova, N., and Lindow, S. E. 2005. Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. Proc. Natl. Acad. Sci. U.S.A. 102:11064-11069.

Ferreira, A. O., Myers, C. R., Gordon, J. S., Martin, G. B., Vencato, M., Collmer, A., Wehling, M. D., Alfano, J. R., Moreno-Hagelsieb, G., Lamboy, W. F., DeClerck, G., Schneider, D. J., and Cartinhour, S. W. 2006. Whole-genome expression profiling defines the HrpL regulon of *Pseudomonas syringae* pv. *tomato* DC3000, allows de novo reconstruction of the Hrp cis clement, and identifies novel coregulated genes. Mol. Plant-Microbe Interact. 19:1167-1179.

Fourie, D. 1998. Characterization of halo blight races on dry beans in South Africa. Plant Dis. 82:307-310.

Fouts, D. E., Abramovitch, R. B., Alfano, J. R., Baldo, A. M., Buell, C. R., Cartinhour, S., Chatterjee, A. K., D'Ascenzo, M., Gwinn, M. L., Lazarowitz, S. G., Lin, N. C., Martin, G. B., Rehm, A. H., Schneider, D. J., van Dijk, K., Tang, X., and Collmer, A. 2002. Genomewide identification of *Pseudomonas syringae* pv. *tomato* DC3000 promoters controlled by the HrpL alternative sigma factor. Proc. Natl. Acad. Sci. U.S.A. 99:2275-2280.

Fu, Z. Q., Guo, M., Jeong, B. R., Tian, F., Elthon, T. E., Cerny, R. L., Staiger, D., and Alfano, J. R. 2007. A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity. Nature 447:284-288.

Gal, M., Preston, G. M., Massey, R. C., Spiers, A. J., and Rainey, P. B. 2003. Genes encoding a cellulosic polymer contribute toward the ecological success of *Pseudomonas fluorescens* SBW25 on plant surfaces. Mol. Ecol. 12:3109-3121.

Gal-Mor, O., and Finlay, B. B. 2006. Pathogenicity islands: A molecular toolbox for bacterial virulence. Cell. Microbiol. 8:1707-1719.

Greenberg, J. T., and Vinatzer, B. A. 2003. Identifying type III effectors of plant pathogens and analyzing their interaction with plant cells. Curr. Opin. Microbiol. 6:20-28.

Grenier, A.-M., Duport, G., Pages, S., Condemine, G., and Rahbe, Y. 2006. The phytopathogen *Dickeya dadantii* (*Erwinia chrysanthemi* 3937) is a pathogen of the pea aphid. Appl. Environ. Microbiol. 72:1956-1965.

Guttman, D. S., Vinatzer, B. A., Sarkar, S. F., Ranall, M. V., Kettler, G., and Greenberg, J. T. 2002. A functional screen for the type III Hrp secretome of the plant pathogen *Pseudomonas syringae*. Science 295:1722-1726.

Hacker, J., and Kaper, J. B. 2000. Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. 54:641-679.

Hirano, S. S., and Upper, C. D. 2000. Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae*-a pathogen, ice nucleus, and epiphyte. Microbiol. Mol. Biol. Rev. 64:624-653.

Hsiao, W. W. L., Ung, K., Aeschliman, D., Bryan, J., Finlay, B. B., and Brinkman, F. S. L. 2005. Evidence of a large novel gene pool associated with prokaryotic genomic islands. PLoS Genetics 1:e62. Published online.

Jackson, R. W., Mansfield, J. W., Arnold, D. L., Sesma, A., Paynter, C. D., Murillo, J., Taylor, J. D., and Vivian, A. 2000. Excision from tRNA genes of a large chromosomal region, carrying *avrPphB*, associated with race change in the bean pathogen, *Pseudomonas syringae* pv. *phaseolicola*. Mol. Microbiol. 38:186-197.

Jelenska, J., Yao, N., Vinatzer, B. A., Wright, C. M., Brodsky, J. L., and Greenberg, J. T. 2007. A J domain virulence effector of *Pseudomonas syringae* remodels host chloroplasts and suppresses defenses. Curr. Biol. 17:499-508.

Joardar, V., Lindeberg, M., Jackson, R. W., Selengut, J., Dodson, R., Brinkac, L. M., Daugherty, S. C., Deboy, R., Durkin, A. S., Giglio, M. G., Madupu, R., Nelson, W. C., Rosovitz, M. J., Sullivan, S., Crabtree, J., Creasy, T., Davidsen, T., Haft, D. H., Zafar, N., Zhou, L., Halpin, R., Holley, T., Khouri, H., Feldblyum, T., White, O., Fraser, C. M., Chatterjee, A. K., Cartinhour, S., Schneider, D. J., Mansfield, J., Collmer, A., and Buell, C. R. 2005a. Whole-genome sequence analysis of *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among pathovars in genes involved in virulence and transposition. J. Bacteriol. 187:6488-6498.

Joardar, V., Lindeberg, M., Schneider, D. J., Collmer, A., and Buell, C. R. 2005b. Lineage-specific regions in *Pseudomonas syringae* pv. *tomato* DC3000. Mol. Plant Pathol. 6:53-64.

Jones, A. M., Lindow, S. E., and Wildermuth, M. C. 2007. Salicylic acid, yersiniabactin, and pyoverdin production by the model phytopathogen *Pseudomonas syringae* pv. *tomato* DC3000: Synthesis, regulation, and

impact on tomato and *Arabidopsis* host plants. J. Bacteriol. 189:6773-6786.

Karamanoli, K., and Lindow, S. E. 2006. Disruption of *N*-acyl homoserine lactone-mediated cell signaling and iron acquisition in epiphytic bacteria by leaf surface compounds. Appl. Environ. Microbiol. 72:7678-7686.

Karlin, S. 2001. Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes. Trends Microbiol. 9:335-343.

Kloek, A. P., Verbsky, M. L., Sharma, S. B., Schoelz, J. E., Vogel, J., Klessig, D. F., and Kunkel, B. N. 2001. Resistance to *Pseudomonas syringae* conferred by an *Arabidopsis thaliana* coronatine-insensitive *coi1* mutation occurs through two distinct mechanisms. Plant J. 26:509-522.

Kvitko, B. H., Ramos, A. R., Morello, J. E., Oh, H.-S., and Collmer, A. 2007. Identification of harpins in *Pseudomonas syringae* pv. *tomato* DC3000, which are functionally similar to HrpK1 in promoting translocation of type III secretion system effectors. J. Bacteriol. 189:8059-8072.

Lamark, T., Kaasen, I., Eshoo, M. W., Falkenberg, P., McDougall, J., and Strom, A. R. 1991. DNA sequence and analysis of the *bet* genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. Mol. Microbiol. 5:1049-1064.

Lan, L., Deng, X., Zhou, J., and Tang, X. 2006. Genome-wide gene expression analysis of *Pseudomonas syringae* pv. *tomato* DC3000 reveals overlapping and distinct pathways regulated by *hrpL* and *hrpRS*. Mol. Plant-Microbe Interact. 19:976-987.

Landfald, B., and Strom, A. R. 1986. Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*. J. Bacteriol. 165:849-855.

Laue, H., Schenk, A., Li, H., Lambertsen, L., Neu, T. R., Molin, S., and Ullrich, M. S. 2006. Contribution of alginate and levan production to biofilm formation by *Pseudomonas syringae*. Microbiology 152:2909-2918.

Lavin, J. L., Kiil, K., Resano, O., Ussery, D. W., and Oguiza, J. A. 2007. Comparative genomic analysis of two-component regulatory proteins in *Pseudomonas syringae*. BMC Genomics 8:397. Published online.

Lawrence, J. G., and Ochman, H. 1997. Amelioration of Bacterial Genomes: Rates of Change and Exchange. J. Mol. Evol. 44:383-397.

Liehl, P., Blight, M., Vodovar, N., Boccard, F., and Lemaitre, B. 2006. Prevalence of local immune response against oral infection in a *Drosophila/Pseudomonas* infection model. PLoS Pathogens 2:e56. Published online.

Lin, N. C., and Martin, G. B. 2007. Pto- and Prf-mediated recognition of AvrPto and AvrPtoB restricts the ability of diverse *Pseudomonas syringae* pathovars to infect tomato. Mol. Plant-Microbe Interact. 20:806-815.

Lindeberg, M., Cartinhour, S., Myers, C. R., Schechter, L. M., Schneider, D. J., and Collmer, A. 2006. Closing the circle on the discovery of genes encoding Hrp regulon members and type III secretion system effectors in the genomes of three model *Pseudomonas syringae* strains. Mol. Plant-Microbe Interact. 19:1151-1158.

Lopez-Solanilla, E., Bronstein, P. A., Schneider, A. R., and Collmer, A. 2004. HopPtoN is a *Pseudomonas syringae* Hrp type III secretion system cyseine protease effector that suppresses pathogen-induced necrosis associated with both compatible and incompatible plant interactions. Mol. Microbiol. 54:353-365.

Ma, L., Jackson, K. D., Landry, R. M., Parsek, M. R., and Wozniak, D. J. 2006. Analysis of *Pseudomonas aeruginosa* conditional Psl variants reveals roles for the Psl polysaccharide in adhesion and maintaining biofilm structure postattachment. J. Bacteriol. 188:8213-8221.

Maillard, J., Spronk, C. A., Buchanan, G., Lyall, V., Richardson, D. J., Palmer, T., Vuister, G. W., and Sargent, F. 2007. Structural diversity in twin-arginine signal peptide-binding proteins. Proc. Natl. Acad. Sci. U.S.A. 104:15641-15646.

Marco, M. L., Legac, J., and Lindow, S. E. 2005. *Pseudomonas syringae* genes induced during colonization of leaf surfaces. Environ. Microbiol. 7:1379-1391.

Melotto, M., Underwood, W., Koczan, J., Nomura, K., and He, S. Y. 2006. Plant stomata function in innate immunity against bacterial invasion. Cell 126:969-980.

Miller, K. J., and Wood, J. M. 1996. Osmoadaptation by rhizosphere bacteria. Ann. Rev. Microbiol. 50:101-136.

Mole, B. M., Baltrus, D. A., Dangl, J. L., and Grant, S. R. 2007. Global virulence regulation networks in phytopathogenic bacteria. Trends Microbiol. 15:363-371.

Monier, J. M., and Lindow, S. E. 2003. Differential survival of solitary and aggregated bacterial cells promotes aggregate formation on leaf surfaces. Proc. Natl. Acad. Sci. U.S.A. 100:15977-15982.

Mougous, J. D., Cuff, M. E., Raunser, S., Shen, A., Zhou, M., Gifford, C. A., Goodman, A. L., Joachimiak, G., Ordonez, C. L., Lory, S., Walz, T., Joachimiak, A., and Mekalanos, J. J. 2006. A virulence locus of *Pseu-*

domonas aeruginosa encodes a protein secretion apparatus. Science 312:1526-1530.

Nomura, K., Melotto, M., and He, S. Y. 2005. Suppression of host defense in compatible plant-*Pseudomonas syringae* interactions. Curr. Opin. Plant Biol. 8:361-368.

Oh, C.-S., Kim, J. F., and Beer, S. V. 2005. The Hrp pathogenicity island of *Erwinia amylovora* and identification of three novel genes required for systemic infection. Mol. Plant Pathol. 6:125-138.

Oh, H.-S., Kvitko, B. H., Morello, J. E., and Collmer, A. 2007. *Pseudomonas syringae* lytic transglycosylases co-regulated with the type III secretion system contribute to the translocation of effector proteins into plant cells. J. Bacteriol. 189:8277-8289.

Paulsen, I. T., Banerjei, L., Myers, G. S. A., Nelson, K. E., Seshadri, R., Read, T. D., Fouts, D. E., Eisen, J. A., Gill, S. R., Heidelberg, J. F., Tettelin, H., Dodson, R. J., Umayam, L., Brinkac, L., Beanan, M., Daugherty, S., DeBoy, R. T., Durkin, S., Kolonay, J., Madupu, R., Nelson, W., Vamathevan, J., Tran, B., Upton, J., Hansen, T., Shetty, J., Khouri, H., Utterback, T., Radune, D., Ketchum, K. A., Dougherty, B. A., and Fraser, C. M. 2003. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. Science 299:2071-2074.

Paulsen, I. T., Press, C. M., Ravel, J., Kobayashi, D. Y., Myers, G. S., Mavrodi, D. V., DeBoy, R. T., Seshadri, R., Ren, Q., Madupu, R., Dodson, R. J., Durkin, A. S., Brinkac, L. M., Daugherty, S. C., Sullivan, S. A., Rosovitz, M. J., Gwinn, M. L., Zhou, L., Schneider, D. J., Cartinhour, S. W., Nelson, W. C., Weidman, J., Watkins, K., Tran, K., Khouri, H., Pierson, E. A., Pierson, L. S., 3rd, Thomashow, L. S., and Loper, J. E. 2005. Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. Nat. Biotechnol. 23:873-878.

Petnicki-Ocwieja, T., Schneider, D. J., Tam, V. C., Chancey, S. T., Shan, L., Jamir, Y., Schechter, L. M., Janes, M. D., Buell, C. R., Tang, X., Collmer, A., and Alfano, J. R. 2002. Genomewide identification of proteins secreted by the Hrp type III protein secretion system of *Pseudomonas syringae* pv. *tomato* DC3000. Proc. Natl. Acad. Sci. U.S.A. 99:7652-7657.

Petnicki-Ocwieja, T., van Dijk, K., and Alfano, J. R. 2005. The *hrpK* operon of *Pseudomonas syringae* pv. *tomato* DC3000 encodes two proteins secreted by the type III Hrp protein secretion system: HopB1 and HrpK, a putative type III translocator. J. Bacteriol. 187:649-663.

Pinheiro, V. B., and Ellar, D. J. 2007. Expression and insecticidal activity of *Yersinia pseudotuberculosis* and *Photorhabdus luminescens* toxin complex proteins. Cell. Microbiol. 9:2372-2380.

Pitman, A. R., Jackson, R. W., Mansfield, J. W., Kaitell, V., Thwaites, R., and Arnold, D. L. 2005. Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. Curr. Biol. 15:2230-2235.

Pritchard, L., White, J. A., Birch, P. R. J., and Toth, I. K. 2006. Genome-Diagram: A python package for the visualization of large-scale genomic data. Bioinformatics 22:616-617.

Pukatzki, S., Ma, A. T., Revel, A. T., Sturtevant, D., and Mekalanos, J. J. 2007. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. Proc. Natl. Acad. Sci. U.S.A. 104:15508-15513.

Pukatzki, S., Ma, A. T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W. C., Heidelberg, J. F., and Mekalanos, J. J. 2006. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. Proc. Natl. Acad. Sci. U.S.A. 103:1528-1533.

Quinones, B., Dulla, G., and Lindow, S. E. 2005. Quorum sensing regulates exopolysaccharide production, motility, and virulence in *Pseudomonas syringae*. Mol. Plant-Microbe Interact. 18:682-693.

Quinones, B., Pujol, C. J., and Lindow, S. E. 2004. Regulation of AHL production and its contribution to epiphytic fitness in *Pseudomonas syringae*. Mol. Plant-Microbe Interact. 17:521-531.

Raaijmakers, J. M., de Bruijn, I., and de Kock, M. J. D. 2006. Cyclic lipopeptide production by plant-associated *Pseudomonas* spp.: Diversity, activity, biosynthesis, and regulation. Mol. Plant-Microbe Interact. 19:699-710.

Robert-Seilaniantz, A., Shan, L., Zhou, J.-M., and Tang, X. 2006. The *Pseudomonas syringae* pv. *tomato* DC3000 Type III effector HopF2 has a putative myristolation site required for its avirulence and virulence functions. Mol. Plant-Microbe Interact. 19:130-138.

Rohmer, L., Guttman, D. S., and Dangl, J. L. 2004. Diverse evolutionary mechanisms shape the type III effector virulence factor repertoire in the plant pathogen *Pseudomonas syringae*. Genetics 167:1341-1360.

Roine, E., Raineri, D. M., Romantschuk, M., Wilson, M., and Nunn, D. N. 1998. Characterization of type IV pilus genes in *Pseudomonas syringae* pv. *tomato* DC3000. Mol. Plant-Microbe Interact. 11:1048-1056.

Rosebrock, T. R., Zeng, L., Brady, J. J., Abramovitch, R. B., Xiao, F., and Martin, G. B. 2007. A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. Nature 448:370-374.

Sarkar, S. F., and Guttman, D. S. 2004. Evolution of the core genome of

Pseudomonas syringae, a highly clonal, endemic plant pathogen. Appl. Environ. Microbiol. 70:1999-2012.

Sawada, H., Kanaya, S., Tsuda, M., Suzuki, F., Azegami, K., and Saitou, N., 2002. A phylogenomic study of the OCTase genes in *Pseudomonas syringae* pathovars: The horizontal transfer of the argK-tox cluster and the evolutionary history of OCTase genes on their genomes. J. Mol. Evol. 54:437-457.

Schell, M. A., Ulrich, R. L., Ribot, W. J., Brueggemann, E. E., Hines, H. B., Chen, D., Lipscomb, L., Kim, H. S., Mrazek, J., Nierman, W. C., and Deshazer, D. 2007. Type VI secretion is a major virulence determinant in *Burkholderia mallei*. Mol. Microbiol. 64:1466-1485.

Sonnhammer, E. L., and Durbin, R. 1994. A workbench for large-scale sequence homology analysis. Comput. Appl. Biosci. 10:301-307.

Spiers, A. J., Bohannon, J., Gehrig, S. M., and Rainey, P. B. 2003. Biofilm formation at the air-liquid interface by the *Pseudomonas fluorescens* SBW25 wrinkly spreader requires an acetylated form of cellulose. Mol. Microbiol. 50:15-27.

Sundin, G. W., and Murillo, J. 1999. Functional analysis of the *Pseudomonas syringae rulAB* determinant in tolerance to ultraviolet B 290-320 nm radiation and distribution of *rulAB* among *P-syringae* pathovars. Environ. Microbiol. 1:75-87.

Tobes, R., and Pareja, E. 2005. Repetitive extragenic palindromic sequences in the *Pseudomonas syringae* pv. *tomato* DC3000 genome: Extragenic signals for genome reannotation. Res. Microbiol. 156:424-433.

Tobes, R., and Pareja, E. 2006. Bacterial repetitive extragenic palindromic sequences are DNA targets for Insertion Sequence elements. BMC Genomics 7:62. Published online.

Toth, I. K., Pritchard, L., and Birch, P. R. J. 2006. Comparative genomics reveals what makes an enterobacterial plant pathogen. Ann. Rev. Phytopath. 44:305-336.

Ude, S., Arnold, D. L., Moon, C. D., Timms-Wilson, T., and Spiers, A. J. 2006. Biofilm formation and cellulose expression among diverse environmental *Pseudomonas* isolates. Environ. Microbiol. 8:1997-2011.

Underwood, W., Melotto, M., and He, S. Y. 2007. Role of plant stomata in bacterial invasion. Cell. Microbiol. 9:1621-1629.

Vencato, M., Tian, F., Alfano, J. R., Buell, C. R., Cartinhour, S., DeClerck, G. A., Guttman, D. S., Stavrinides, J., Joardar, V., Lindeberg, M., Bronstein, P. A., Mansfield, J. W., Myers, C. R., Collmer, A., and Schneider, D. J. 2006. Bioinformatics-enabled identification of the HrpL regulon and type III secretion system effector proteins of *Pseudomonas syringae* pv. *phaseolicola* 1448A. Mol. Plant-Microbe Interact. 19:1193-1206.

Ventre, I., Goodman, A. L., Vallet-Gely, I., Vasseur, P., Soscia, C., Molin, S., Bleves, S., Lazdunski, A., Lory, S., and Filloux, A. 2006. Multiple sensors control reciprocal expression of *Pseudomonas aeruginosa* regulatory RNA and virulence genes. Proc. Natl. Acad. Sci. U.S.A. 103:171-176.

Vernikos, G. S. and Parkhill, J. 2006. Interpolated variable order motifs for identification of horizontally acquired DNA: Revisiting the *Salmonella* pathogenicity islands. Bioinformatics 22:2196-2203.

Vieira, J., Mendes, M. V., Albuquerque, P., Morades-Ferreira, P., and Tavares, F. 2007. A novel approach for the identification of bacterial taxa-specific molecular markers. Lett. App. Microbiol. 44:506-512.

Vigneux, F., Zumbihl, R., Jubelin, G., Ribeiro, C., Poncet, J., Baghdiguian, S., Givaudan, A., and Brehelin, M. 2007. The *xaxAB* genes encoding a new apoptotic toxin from the insect pathogen *Xenorhabdus nematophila* are present in plant and human pathogens. J. Biol. Chem. 282:9571-9580.

Vodovar, N., Vallenet, D., Cruveiller, S., Rouy, Z., Barbe, V., Acosta, C., Cattolico, L., Jubin, C., Lajus, A., Segurens, B., Vacherie, B., Wincker, P., Weissenbach, J., Lemaitre, B., Medigue, C., and Boccard, F. 2006. Complete genome sequence of the entomopathogenic and metabolically versatile soil bacterium *Pseudomonas entomophila*. Nat. Biotechnol. 24:673-679.

Waterfield, N. R., Bowen, D. J., Fetherston, J. D., Perry, R. D., and ffrench-Constant, R. H. 2001. The tc genes of *Photorhabdus*: A growing family. Trends Microbiol. 9:185-191.

Wei, C. F., Kvitko, B. H., Shimizu, R., Crabill, E., Alfano, J. R., Lin, N. C., Martin, G. B., Huang, H. C., and Collmer, A. 2007. A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. Plant J. 51:32-46.

Welch, R. A., Burland, V., Plunkett, G., III, Redford, P., Roesch, P., Rasko, D., Buckles, E. L., Liou, S. R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G. F., Rose, D. J., Zhou, S., Schwartz, D. C., Perna, N. T., Mobley, H. L. T., Donnenberg, M. S., and Blattner, F. R. 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024.

Wood, J. M. 1999. Osmosensing by bacteria: signals and membrane-based sensors. Microbiol. Mol. Biol. Rev. 63:230-262.

Xing, W., Zou, Y., Liu, Q., Liu, J., Luo, X., Huang, Q., Chen, S., Zhu, L., Bi, R., Hao, Q., Wu, J.-W., Zhou, J.-M., and Chai, J. 2007. The structural basis for activation of plant immunity by bacterial effector protein AvrPto. Nature 449:243-247.

Yu, J., Penaloza-Vazquez, A., Chakrabarty, A. M., and Bender, C. L. 1999. Involvement of the exopolysaccharide alginate in the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. Mol. Microbiol. 33:712-720.

Zhao, Y., Thilmony, R., Bender, C. L., Schaller, A., He, S. Y., and Howe, G. A. 2003. Virulence systems of *Pseudomonas syringae* pv. *tomato* promote bacterial speck disease in tomato by targeting the jasmonate signaling pathway. Plant J. 36:485-99.

AUTHOR-RECOMMENDED INTERNET RESOURCE

Pseudomonas-Plant Interaction website: www.pseudomonas-syringae.org